REPORT RESUMES

ED 013 045 08 FL 000 518

STANDARDIZATION OF A SPANISH LANGUAGE ADULT INTELLIGENCE SCALE. FINAL REPORT.

BY- GREEN, RUSSEL F. MARTINEZ, JUAN N.
PUERTO RICO UNIV., RIO PIEDRAS
REPORT NUMBER BR-5-0081 PUB DATE MAR 67
CONTRACT OEC-3-10-128
EDRS PRICE MF-$1.25 HC-$11.64 291P.

DESCRIPTORS- *ADULT EDUCATION, *INTELLIGENCE TESTS, *LANGUAGE RESEARCH, *SPANISH, *TEST CONSTRUCTION, STATISTICAL DATA, TEST RELIABILITY, PREDICTIVE VALIDITY, ITEM ANALYSIS, TESTING PROGRAMS, RESEARCH CRITERIA, PUERTO RICO, PROJECTO EIWA, WAIS PROJECT, WECHSLER ADULT INTELLIGENCE SCALE,

A NEED FOR AN ADEQUATELY DEVELOPED AND STANDARDIZED INTELLIGENCE SCALE IN THE WHOLE OF LATIN AMERICA RESULTED IN THE WAIS PROJECT WHOSE AIM WAS TO TRANSLATE INTO SPANISH, ADAPT TO SPANISH CULTURE, AND STANDARDIZE THE WECHSLER ADULT INTELLIGENCE SCALE IN PUERTO RICO. FOLLOWING A DISCUSSION OF THE FOUR GENERAL GOALS, THE PROJECT REPORT OUTLINES THE NINE PHASES BY WHICH THE GOALS WERE ACCOMPLISHED. IN THE SECTION THAT FOLLOWS, THE RESULTS OF THE STATISTICAL ANALYSES, SUPPORTED BY NUMEROUS TABLES, ARE PRESENTED IN DETAIL. DATA COVER (1) ANALYSES OF THE TOTAL POPULATION REPRESENTED, (2) THE GUATEMALA EXTENSION OF THE PROJECT, (3) THE DEVELOPMENT OF THE NEW MANUAL, (4) THE STANDARDIZATION SAMPLE OF THE TOTAL POPULATION REPRESENTED, (5) THE PROBLEM OF PREDICTING THE RATE OF CHANGE OF IQ AND THE NEED FOR AUTOMATIC ADJUSTMENT OF NORMS OR RESTANDARDIZATION, AND (6) THE PREDICTION OF NORMS FROM A RELATIVELY SMALL SAMPLE. A DISCUSSION OF THE SERIOUS QUESTIONS THAT CAN BE RAISED ABOUT THE QUALITY OF THE TEST PRECEDES A SECTION CONTAINING CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS. EXTENSIVE APPENDIXES INCLUDE A SPANISH TRANSLATION OF THE TEST MANUAL AND A PAPER ON RANDOM SAMPLING FROM TOTAL POPULATIONS. (AB)

ED013045

FINAL REPORT
PROJECT NO. 1963
CONTRACT NO. O. E. 3-10-128

STANDARDIZATION OF A SPANISH LANGUAGE ADULT INTELLIGENCE SCALE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

MARCH 1967

FL 000 518

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
OFFICE OF EDUCATION
BUREAU OF RESEARCH

FINAL REPORT
PROJECT NO. 1963
CONTRACT NO. O. E. 3-10-128

STANDARDIZATION OF A SPANISH LANGUAGE ADULT INTELLIGENCE SCALE

MARCH 1967

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

OFFICE OF EDUCATION
BUREAU OF RESEARCH

STANDARDIZATION OF A SPANISH LANGUAGE ADULT INTELLIGENCE SCALE

Project No. 1963
Contract No. O. E. 3-10-128

Russel F. Green

Juan N. Martinez

March 1967

The research reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

The University of Puerto Rico
Rio Piedras, Puerto Rico

TABLE OF CONTENTS

Page

Acknowledgements ii

I Introduction 1

II Method 6

III Results 15

IV Discussion 77

V Conclusions, Implications and Recommendations . . . 83

VI Summary 88

Bibliography 98

Appendixes

I Test Manual

II Random Sampling from Total Populations
A Demonstration of its Power and Flexibility

ACKNOWLEDGMENTS

It is difficult to be sure that all who have made significant contributions to a project such as this one have received full credit for their efforts. We would like to list the major contributors and to add that without the efforts of many unnamed individuals the project still could not have been completed.

1. Russel F. Green, Ph.D., Director
   University of Puerto Rico and University of Rochester
2. Juan Nicolás Martínez, Ph.D., Co-Director
   University of Puerto Rico
3. Mrs. Elia Hidalgo de Christensen , Assistant Director, June 1963 - June 1965
4. Miss Laura Leticia Herrans, Research Associate and then Assistant Director, July 1965 - June 1966
5. Mrs. Fredeswinda de Jesús, Testing Supervisor and then Assistant Director, July 1966 - March 1967
6. Mr. Edwin B. Fernández, Research Associate
7. Mr. Gregorio López, Research Associate
8. Miss Jane Podstawski, Research Associate
9. Mrs. Nora Rivera de Vázquez, Research Assistant

There were many others, including some who worked for less time as research associates, the valuable help from administrative assistants, testers who worked part time and in some cases full time, research assistants, and student assistants. The complete count surely run to more than 100 persons.

It should be emphasized that all of the nine persons named above made significant contributions to the writing of the manual which appears as Appendix I of this report. The group who assumed the primary responsibility for the final draft of the Spanish manual included Mrs. Fredeswinda de Jesús, Mr. Edwin B. Fernández, Mrs. Nora Rivera de Vázquez, Dr. Juan Nicolás Martínez, and Dr. Russel F. Green. The manual, then, is the product of many authors who have worked long and cooperatively in its production.

The contributions of all who have worked so long, faithfully, and cooperatively are acknowledge and appreciated. All are in a real and important way co-authors of this project.

RUSSEL F. GREEN

JUAN N. MARTINEZ

I. INTRODUCTION

A. Background

As is now generally known, the WAIS project (Proyecto EIWA) is a project to translate into Spanish, adapt to Spanish culture, and to standardize the Wechsler Adult Intelligence Scale in Puerto Rico. Actually, the charge to the project from the principal supporting agency (The U. S. Office of Education) was to use a broad definition of Spanish culture and to avoid local colloquialisms in order that the resulting test could be more readily standardized in other countries. Norms for other Spanish speaking will, of course, vary somewhat because the use of appropriate colloquialisms in answers must receive full credit and because of variations in local history, population composition, educational levels, etc.

Little need be said about the need for the project. The need was fully attested to by the fact that, because nothing better was available, numerous psychologists and agencies throughout Latin America have been using inadequately developed forms with inappropriate norms for many years. Insofar as could be determined, this is the first instance of an adequately developed and standardized intelligence scale in the whole of Latin America.

One should immediately agree that the use of improper norms is not in itself necessarily bad. The translation in use might still have predictive efficiency equal to that of the original. Satisfactory evidence on this point does not seem to have been developed. More important, however, is the fact that this practice almost surely leads to incorrect labeling of, and therefore improper attitudes toward, individuals. Such labeling and attitudes can have profound effects on the individuals all their lives. Note carefully that mislabeling and its subsequent effects can of itself produce a spurious validity. It is spurious in that it is due to effects that have nothing to do with what a person could have done but rather have to do with what he was induced to do. For this reason, if for no other, a properly developed and standardized test was sorely needed.

B. Objectives

Let us proceed now to a brief discussion of the over-all goals of the project. This will be followed by an extensive report of the methods used, the results achieved, and finally a post hoc evaluation of the appropriateness of the project.

There were four general goals that it was anticipated that this project would achieved. Let us list them in the time sequence in which we expect that they would be achieved. The time sequence is also to some degree an indication of priorities.

1. To produce and standardize in Puerto Rico an individual intelligence scale written in the Spanish language and adapted to Spanish culture for the age range 16 through 64.
2. To develop a test and a procedure which will make it possible to adapt the test to other Spanish American countries with a minimum of cost and effort.
3. To develop information which will allow us to set up an automatic correction factor which will adequately account for much of the time related shift in mean score that is normal for tests of this kind.
4. To incorporate into the overall effort the investigation of a variety of questions of general interest to Psychology.

The first three of these four goals will be discussed and clarified in turn. The fourth is listed at this time merely to report that it was a real and important element in the overall planning and conduct of the project.

1. The production of an individual intelligence scale for native Spanish speaking adult Puerto Ricans

The need and benefits from such a scale are well known. Yet, so far as we have been able to discover, no such scale existed in the whole of Spanish America. There are quite a number of translations and partial adaptations of existing scales that are being used in various countries but completely adequate norms apparently do not exist for any of these. Their use, then, is bound to be to some degree unsatisfactory. Test materials used will often be inappropriate. Errors of measurement are likely to be larger than normal. Reference statistics such as the population mean and variability are not known. Short comings such as these increase subjectivity and guesswork in interpretations of thest results.

By permission of the Psychological Corporation in New York City, the project undertook the task of adapting the Wechsler

Adult Intelligence Scale (WAIS) for use among Spanish speaking adults. There are those who feel that developing a test for a given cultural group should be an entirely new and independent effort. We believe that there is much advantage to starting with an existing test - providing the owners of the test are willing to have the test modified to whatever extent required to make it completely suitable for use within the culture to which it is to be applied. The Psychological Corporation has been most cooperative and agreeable in this regard.

The advantages of adapting existing tests are real and important. One starts with a known model to follow. There will have accumulated large amounts of information (especially concerning validity) and experience from use of the original test, much of which will prove to be valid in the new cultural setting. General criteria for scoring are available and can be used as a guide in setting up scoring keys for the new test. In short, very large amounts of work need not be repeated. Disadvantages accrue only if the owners of the test interfere with or unnecessarily restrict efforts at adapting the materials to the new population or cultural group.

All these observations are fully supported by the successful adaptation of the WAIS in several countries in Europe, Africa, and the Far East.

2. So far, as we know, there is no individual adult intelligence scale in the Spanish Language which has been completely adapted and standardized in any of the 20 or so Spanish speaking political units in the Western Hemisphere.

It is desirable, therefore, to use test materials which will be as generally applicable to Spanish American subcultures as can be. This we have tried to do. This effort accounts for our having sent the material to other people in other countries for review.

No one has even written a truly culturally free test - especially one as complex as this. This being the case any other Spanish speaking country that wishes to use the EIWA (Spanish WAIS) should carry out a restandardization to insure that scoring guides and norms are appropriate for that particular country. To retest up to 2,000 persons in each of about 20 countries would be a very expensive and arduous task. Can it be made more practical to properly adapt norms than this?

We had hoped to investigate this problem very carefully. The plan was this. An initial stratified random sample of 250 persons was to be used to predict what the norms from a large standardization sample of 1,800 would eventually be. If we, could successfully predict the large sample norms from the small, then we would have established a justification for others to adapt the test to their countries with a small sample of about 250. This will be illustrated later, the plan was only partly realized.

3. Experience has shown that the average IQ score increases steadily across time in any country that is advancing educationally and technologically.

In the United States, restandardizing is needed each 12 to 18 years. Puerto Rico is changing rapidly. The average number of years of schooling is increasing. Mass communication is improving. A large middle class has formed. And still other changes could be specified.

How do we approach the problem of avoiding early obsolescence of the test we will have produced? There are various possibilities. We approached the problem in the following manner. We used parent-child differences as a partial means of developing an index of the rate of increase of raw scores in the population. For each member of a randomly selected portion of the group of subjects who had children and were in the 35 to 64 age range, we also tested one of their children who was between 16 and 25. Conversely for each member of a randomly selected portion of the subjects who were in the 16 to 19 age range we tested one of their parents. Since the subject drawn for the standardization sample will be stratified random samples of their respective age brackets, the parents and children of these subjects should also be approximately stratified random samples of their respective age ranges.

Since parents and children are genetically related, and since we were using unbiased samples of each, we could expect that there would be no genetic basis for differences between the average ability levels of parents and their children. Most of the mean difference between parents and children which would be observed would be due to the impact of the changing culture on those abilities which are important to earning scores on the test. Such changes will accrue from such factors as the changing number of years of schooling people are getting, television, increases in the reading of newspapers, and books. It was recognized that some small portion of any differences observed might be due to a slow decline in mean raw scores from early adulthood on, as noted by Wechsler (3) but to the extent that such decline occurs it would merely add to differences due to other cultural factors which might be present.

A further look at the impact of cultural factors other than the simple number of years of schooling was to be obtained by collecting three special subgroups of ages 25-29, 40-49, and 55-64 who were matched in terms of number of years of schooling.

These two sources of information were used to estimate the rate of changes of IQ in the Puerto Rican population. From this estimate a correction factor based on the number of years since the test was constructed was developed. It should be possible to keep the mean score fairly accurate for perhaps 15 to 20 years through such a device. Eventually, however, restandardization will be necessary. This fact makes a goal number three important even to Puerto Rico.

This, then, represents an overview of the project insofar as goals 1, 2, and 3 are related to the needs of Puerto Rico in particular and Spanish America in general. Goal 4 is more remote and need not be discussed in detail now. There will, in time, be additional special analyses made of portions of the data collected. Since each day has only 24 hours, these analyses will have to come after the first three goals are reached and the project is officially concluded.

There were two more specific goals included in the project. The first of these was to attempt to provide some evidence of concurrent validity for the scale being constructed. To do this it was, in effect, necessary to standardize a second test as well.

The second more specific goal was to provide for a short form of the WAIS. This involved investigating which combination of, say four subtests would be in some sense the best prediction of full scale performance.

II. METHOD

No really clear distinction between method and results can be made. In this section a relatively brief outline summary of the methods used will be presented. The project proceeded through some nine relatively distinct phases. Each of these will be outlined in turn in this section. The results of the statistical analyses involved in the four major data processing phases will be presented in detail in the next section.

A. Translation

Since the Wechsler Adult Intelligence Scale has been adopted as the model, the first step was to make a direct translation of the WAIS into Spanish. This first step was taken prior to the official beginning of the contract between the Office of Education and the University of Puerto Rico. The initial translation was made by Dr. Carlos Albizu-Miranda, of the Department of Psychology of the University of Puerto Rico. This initial translation was reviewed and revised twice by other members of the same department.[1] Up to this point all persons working on the translation had Spanish as their native language and all also spoke English. These early translators also initiated a study of word frequencies in order to develop a guide as to possible word order in the vocabulary subtest and also in order to develop possible alternate words for words now included in the English WAIS. The whole staff of the Department of Psychology was asked to cooperate in the word list study. Preliminary study of other subtests, -especially of certain of the non-verbal materials was also begun. It was judged that by the time that contracts were signed the primary translation task had been completed.

The contract for the project officially came into being on June 10, 1963. The first meeting of the research team that had been formed occurred on July 1, the day after the director of the project arrived in Puerto Rico.

---

[1] Miss Elia Hidalgo, Miss Laura Leticia Herrans, and Miss Sylvia Echegaray.

B. Adaptation

The initial task of the research team that was formed to carry out the project[2] was to carefully review the translation that had been completed. The native language of the director of the project is English. It is important for a project such as this that at least one member of the team have the language in which the test was originally cast as his native tongue. It is important for this reason. The English speaking person provides more precise interpretation of the English version of the test which the Spanish speaking person can then write into the translation. Since few people achieve complete and equal mastery of two languages, and since no such persons were available to the project, close cooperation of this sort clearly was an essential ingredient.

During this review, all test materials that in any way appeared to be specific to the culture of the United States were eliminated or modified. New test materials which were judged to be pertinent to a culture of Spanish origins were developed and included. Some items were merely modified in a manner that would make them clearly appropriate for Spanish examinees. A summary of the changes in item content is contained in Appendix I, page 6. Actually, much of the new material introduced and much of the old material kept can in one sense be regarded as being relatively free of culture. Attitudes and experiences with any materials may vary from one culture to another but whenever we could we chose kinds of information or behavior which we were confident occur in all cultures. In this sense, it is hoped that the new Spanish version of the WAIS will actually prove to be more culture free than the English version was. This is important to our hope that the test can be economically adapted to each of the various countries in Spanish America.

C. Assembly and Review of Preliminary Manual

The revised and new materials were then assembled into an experimental form of the test. This form was about 25% longer than the final form was expected to be. This experimental version of the test was reviewed by two different types of reviewers. First, we retained two linguists who specialize in the study of regional and national differences in the use of Spanish. They were asked to base their revision on four criteria. First, the Spanish used were to be correct although the emphasis was not to be on producing a highly literary style. Second, the instructions were to be as simple and concise as possible. Third, the parts of the instructions to be read to the subjects were to be expressed in colloquial language which would be as clear, simple and precise as possible. Fourth, that the verbal materials were to be as meaningful and fair as possible for all the subcultures of Spanish America.

---

[2] Dr. Russel F. Green, Director; Dr. Juan N. Martínez, Associate Director (on leave during 1963-1964 academic year); Mrs. Elia Hidalgo de Christensen, Assistant Director; Miss Laura Leticia Herrans, Research Associate; Mrs. Mercedes Pacheco de Izquierdo, Research Associate; and within a short time Mrs. Jane Podstawski, Research Associate.

After revising the manual according to the suggestions of the two linguists, the manual was sent to psychologists in Guatemala, Venezuela, and Uruguay for further reviews. These reviewers were asked to follow much the same criteria as the linguists did and further to keep in mind the fact that the test must be understood in their country. By this it was meant that in some cases the question would be whether the Spanish used is understood in their country rather than whether it is that which is most used. They were also asked to comment on the appropriateness of the test material in their country. The comments and suggestions of these reviewers were collated and still another experimental manual prepared.

D. Training of examiners and Pretesting the instructions and item content

While the review of Psychologists in the other countries was being done, a pretesting of the test materials on small samples was started. This was needed to provide the group with the experience in recruiting subjects, with an opportunity to give training to testers, and with data from which to prepare tentative scoring guides for the Vocabulary, Comprehension, and Analogies subtests. The vocabulary subtest was first pretested by having about 120 high school juniors and seniors write definitions of the words. These written definitions were used to carry out an initial item total-test discrimination analysis and to start a tentative scoring guide for this subtest.

All examiners were given extensive training in the administration of the test. Arrangements were made for every examiner to be observed through one way vision window, in order that his technique, accuracy, diction, rapport with and responsiveness to the subject could be observed and suggestions for improvements made. All examiners both observed such sessions and performed in such sessions. All examiners joined in discussions and evaluations of such sessions.

E. Disparate Group Testing.

Following the initial try-outs and training, all subtests were further pretested by administering them to two groups of subjects (by testing one subject at a time, of course) who were known to be widely different in intellectual ability. The one group consisted of 24 University students and the other group consisted of 20 individuals who had been identified as being in the border line defective to low normal range of intelligence. The performance of these two groups was compared on every item. Any item that failed to differentiate between these two groups was dropped from the test. It was expected, of course, both that some of the old items retained and some of the new materials added would fail to differentiate between intellectual levels and this is why the experimental form of the test was made about 25% longer. It should also be mentioned that our definition of failure to differentiate could not be too stringent for this pretesting since the group was still finding a few changes in instructions, items, and procedures to be required. Small disparate groups comparisons like this can do little more than highlight items that are definitely of little or no value. In a very few instances there were in fact items that failed to differentiate between these groups and so were dropped. It had been expected that the mean of the general population would fall approximately half way between these two groups. Observation of this halfway point indicated that the overall level of the difficulty of the test should be approximately correct.

F. The "254" or item analysis sample.

Since the difficulty level of the test material appeared appropriate, the instructions seemed to be functioning well, and the item content seemed well established, the project next moved into the next phase. This phase involved a relatively major commitment of time and resources. It involved obtaining a representative sample of about 250 persons from the population of Puerto Rico, contacting them, testing them, and subjecting the resulting data to rigorous statistical analyses. Much of the eventual degree of success of the project was to depend on the successful completion of this phase.

For a relatively small sample such as this, it is crucial that a method of selection be available which will clearly and unequivocally produce a representative sample. The method used was a multi-stage, partially urban-rural stratified, random sampling. The sampling methods used were also critical to the eventual goal of making it possible to standardize the test in other countries at the least possible cost. This coupled with the fact that so far as is known this is the first time that random sampling methods have been used to select samples for test standardization, led the Director of the project to write a journal article describing the methods and their results in detail and to submit it for publication. A copy of the article is included as Appendix II of this report. It was submitted to the Psychological Bulletin but as of now no word as to acceptance or rejection has been received. A very brief review of the article on the sampling procedures and their outcome is presented in the next paragraph for the sake of continuity.

Through the cooperation of Dr. Miguel Valencia of the Institute of Statistics at the University of Puerto Rico and of the Bureau of Labor Statistics of the Island Government, a random sample of the island population was obtained. This random sample was subdivided into balanced halves and then one was designated as the "primary" sample on the basis of the flip of a coin. The primary sample consisted of 264 individuals. The secondary sample was used as a source for replacement of people in the primary sample who for some reason could not be tested. Only about three per cent of the people contacted refused to be tested. Nearly 20 per cent, however, had to be substituted from the secondary sample. These replacements were due to deaths, moves, people having left the island, etc. A total of 254 persons were tested. These 254 were compared with the 1960 census in terms of their distributions on age, sex, urban-rural residence, occupation, education, and region of the island. Several combinations of the joint distributions of the control variables were also compared with corresponding census data. The only discrepancy of note was the fact that the sample had a surplus of people reporting 11 and 12 years of schooling relative to the census and a deficiency of people reporting no education. This discrepancy is consistent with the increasing amount of education on the island but probable was somewhat more than would have been predicted from the rate of increase .

The data from these 254 testings were used to obtain estimates of item difficulty, compute item-test correlations, investigate the homogeneity of item content, make special studies of the relative validity of alternate scoring keys for various items, investigate the merits of certain of the time bonuses used in the scoring, and to compare the Puerto Rican test performance characteristics with those found in the States. All these analyses are reported in the Results section of this report.

The critical outcome of the 254 sample analysis was the discovery that the difficulty level of the test was set too high. The population average turned out to be at the level of the low group in the disparate group testing. This fact forced an important modification in the work outline of the project, which was to have important effects on certain goals of the project.

G. The Guatemala extension of the project.

While the testing of the 254 sample was still in progress members of the Department of Psychology of the University of Guatemala asked that they be permitted to test a sample of people in that country. Since this would provide valuable information concerning the usefulness of the test in other Spanish speaking areas, it was considered a highly desirable extension of the project. An application for supplemental funds to cover this extension was approved by the Office of Education.

Due to delays in setting up the contracts and in shipment of data the results of the Guatemala testing were not available until quite some time after the standardization testing had been started. Hence the Guatemala results were not considered in the immediately following phases of the project. Fortunately, the Guatemala results were very encouraging in that the test functioned there even better in certain respects than in Puerto Rico! A summary of these results appears in section 2 of the Results Chapter of this report.

H. The "124" or second item analysis sample.

Since the 254 sample had shown that the scale was too difficult and after consultation with the Psychological Corporation and with the Office of Education, it was decided to initiate a new and unexpected cycle of item development and analysis. The object was to develop new easier item material to enter into seven of the subtests. In order to keep the length of the full scale manageable it was also necessary to drop a few of the most difficult items.

It was further decided to restrict the new sample to the north-east region of the island which includes San Juan and rural areas extending perhaps 15 miles to the west, about 25 miles to the east, and southward approximately it extends to the center of the island. This was done in part for economy reasons.

The economy was needed because going back and undertaking a new item development phase meant that we would have to out back on the size of the standardization sample in order to be able to complete the project within the time and money resources alloted to it. We wished to avoid cutting any more than was absolutely essential.

The results of this phase are presented in the next section of this report. It is sufficient to mention here that this phase was deemed sufficiently successful to warrant continuing to the next phase -- the major standardization effort.

I. The Standardization phase.

The information gained from the previous phases was pooled and used as a guide for the assembling of the standardization manual. The instructions were reviewed and a few minor revisions made. Items were selected to maximize the range across which discrimination could be achieved and to attempt to reduce somewhat the extent of skewness in the resulting raw score distributions. The expected means of the subtests could be predicted quite well and these were set as high as could be and still be consistent with the need for items which would discriminate at the low levels.

Dr. Miguel Valencia of the Institute of Statistics in the College of Commerce of the University of Puerto Rico again cooperated in the selection of samples to be tested. The standardization sample was drawn in two stages -- the first in the spring of 1965 and the second in the summer of 1965. The two halves of the sample compared very well in all respects except one. They were reliably different in urban-rural composition. Fortunately each was off by virtually the same amount in opposite directions so this potential bias cancelled itself. A total of 1176 subjects were drawn for the standardization sample. Some 440 extra cases were drawn for the parent-child and the education balanced groups.

As an aside it should be noted that the Department of Labor had revised their sampling techniques between the time that we drew our 254 sample and the time that we drew our standardization sample. The outcome on the urban-rural percentages which we observed does not lead to confidence that their revision resulted in more exact samples being drawn.

The Standardization sample actually tested consisted of 1127 subjects and was quite representative of all segments of the population. Five tables describing the samples are all presented in Chapter 2 of the test manual which is presented as Appendix I of this report. The information presented in each table is (1) the percent in the 1960 census, (2) the percent in the 1965 sample drawn, and (3) the percent in the sample tested of various combinations of six variables. The six variables are age, sex, urban-rural residence, geographic region, occupation and education.

Inspection of Tables 1 to 5 will reveal the fact that the sample tested is substantially like the sample drawn in all respects. There were, however, some differences between the sample drawn and the 1960 census. The principal difference again involve education and occupation, again involving higher average education and fewer people who are not classifiable either as student, or in the labor force, or as housewives. In this sampling there was also a tendency to get a few too many persons in the 16 to 19 age bracket and too few males in the 20-24 age bracket. The variations in age distribution may be related to migration patterns between Puerto Rico and the United States mainland.

It will be noted that there are a few (17) included in the sample tested who had not reached their 16th birthday. Most of these were within 3 months but unless an error in recording occurred one was only 15 years and no months. This one person might better have been eliminated. The means and variances of these 17 individuals were statistically indistinguishable from the means and variance of the 16-19 age groups, since they were tested as part of the sample they were left in.

Testing was carried out by a total of some 20 part and full time Psychologists each of whom was given special training in the administration of the test. Most, but not all of these Psychologists had earned Masters Degrees. Those who had not Masters Degrees had the equivalence.

As before the tester was given the name, address, sex, age, education, and last known occupation of each person he was to recruit and test. As before the percent of persons who were actually contacted and who refused to cooperate was very low. Refusals occurred 58 times which is a 4% rate. A total replacement rate of 17% was required. At the end of the testing 49 persons drawn for the sample were dropped because of time limitations and because it had become so uneconomic for the testers to attempt to continue. At the end of the testing period all of the testers were being paid by the test. If one wishes to more nearly complete the testing, he should arrange for at least the latter phases of it to be done by full time salaried employees rather than by part time "piece rate" employees. The 49 dropped from the sample included more urban than rural dwellers but since it had been anticipated that some persons would need to be dropped, a surplus of urban cases remaining had been allowed for. This was done by switching about 10 rural lost cases to urban replacements. With this exception all replacements were always as nearly like the lost cases as could be found in the auxiliary samples.

Fifty people drawn for the sample could not be tested because of various illness which included 18 persons identified as mentally ill, 18 persons who had had various physical disabilities such as heart attacks, 8 alcoholics, and 6 severely mentally retardates. Forty-nine persons in the sample could not be traced, 47 had moved to the United States, 13 had simply moved to another town in an area in which testing had already been completed, the native language of 7 was not Spanish, 5 were over 65 years old or more, 2 were dead, and 26 were lost for various other reasons.

When a completed protocol was returned to the project office, it was checked by members of the office staff for completeness. Following this the scoring was checked in its entirety. Following the scoring check, the numbers of the subject and demographic information describing him were coded and entered on data sheets for IBM punching. Individual item responses were also entered on data sheets for IBM punching - one card for each subtest except for Digit Span and Digit Symbol substitution. Scores for these two tests were entered on a 13th card summarizing the subtest scores and also the Verbal Total, Performance Total and Full Scale Total.

The entry of data on the data sheets was checked. The cards were punched by the IBM tabulation center of the University of Puerto Rico. Listings were made of the punched cards and these listings were then checked against the data sheets and in some cases against the protocols. Errors were corrected and then the cards were duplicated.

The duplicate decks were sent to the University of Rochester, Rochester, New York for statistical processing on an IBM 7074 facility under the direct supervision of the director of the project. A special program was prepared for making a final consistency check of the demographic data codes, subtest scores, and Verbal, Performance, and Full Scale Totals on all ten data cards of each subject. Some 75 cards having errors were thus identified and corrected. This is a rate of six tenths of one percent of the cards. Because of this and previous checks, it is unlikely that errors in the data exceed a rate of perhaps one tenth of one percent.

The results of the data analyses are presented in the Result section of this report.

This, then, constitutes a review of methods and phases through which the project went. All this serves as background against which to evaluate the results presented in the next major section of this report.

III. RESULTS

A. Data analyses, sample No. 1, N = 254, Total population represented.

1. Item indices

The results of the analyses of the item data from the first representative sample were all very favorable except for one very important aspect. The test proved to be too difficult for the general population. This result is illustrated in Table 1. Here a set of scaled scores has been set up on each subtest and, for the sake of comparison, this has also been done on each of the unweighted sums of raw scores comprising a verbal total, a performance total, and a full scale score.

Note that of the eleven subtests, none of them scale well at the low end but that Comprehension and Object Assembly do better than most. Note also that the three total scores scale adequately throughout the range. This latter implies that at least a three score test could have been set up without difficulty. It would have almost surely resulted in relatively unreliable low scores but it could have been done.

It is also clear that the test as a whole had a much higher ceiling than was needed for this population. This last statement is illustrated by the substantial ranges of scores that are associated with scaled scores of 19 for many of the subtests as well as for the V, P, and FS scores.

In Table 2, the means and standard deviations of the eleven subtests along with the item difficulty and item subtest correlations for the nine subtests for which this was appropriate have been listed. Only two means are as much as twice their standard deviations. Hence skewing is considerable in most subtests. As a matter of fact, the population average on each subtest was essentially at the level of the low group used in the disparate-group comparison. This, of course, strongly suggests that that group was relatively much better than had been thought. Item difficulties are shown to be consistent with all that has been said. Rarely did we get an item such that virtually all the population got it correct. In spite of this the item versus subtest total correlations (product moment r's corrected for contamination) are generally quite favorable. In fact many of the low correlations are low because of the very unfavorable pass-fail splits that occurred and are at or near their potential maximums.

Table 1

SCALED SCORES FOR ORIGINAL SAMPLE OF 254 SUBJECTS

(The bracketed numbers are scaled scores for the 192 Subjects who had 4 years or more of education)

Scaled Scores	Prop	Arithmetic	Information	Analogies	Comprehension	Digit Span	Vocabulary	Picture Completion	Block Design
19	.999	17	33-36	29-32	33-36	15-17	89-102	27-28	44-48
18	.998	16	32	27-28	32	14	83-88	26	40-43
17	.994	15	31	24-26	30-31	13	75-82	25	38-39
16	.985	14	28-30	22-23	28-29	12	70-74	23-24	36-37
15	.967	13	24-27	20-21	25-27	11	63-69	21-22	31-35
14	.933	11-12	22-23	16-19	23-24	-	57-62	20	29-30
13	.878	10	17-21	13-15	20-22	10	47-56	17-19	25-28
12	.798	9	13-16	10-12	18-19	9	39-46	14-16	19-24
11	.691	8	10-12	7-9	16-17	-	30-33	11-13	17-18
10	.566	7	8-9	6	14-15	8	24-29	9-10	13-16
9	.434	6	6-7	5	11-13	7	19-23	7-8	9-12
8	.309	5	5	4	9-10	-	14-18	5-6	7-8
7	.202	4	4	3 (3)	8	6	8-13	3-4	3-6
6	.122	3	3	2 (2)	6-7	5	3-7	2	1-2 (6)
5	.067		2	1	4-5	4	2	1	0
4	.033	2	1	0 (1)	3	3	1 (3)	0	0
3	.015		0	0 (0)	2	1-2	0 (1)	0	0
2	.006	1	0	0	1	0	0	0	0
1	.002	0	0	0	0	0	0	0	0
0	.001	0	0	0	0	0	0	0	0

Table 1

(Cont.)

Scaled Scores	Picture Arrangement	Object Assembly	Digit Symbol	These scales are directly from unweighted sums of raw scores		
				V	P	FS
19	46-60	45-55	80-100	201-240	202-281	421-521
18	45	43-44	69-79	186-200	199-201	391-420
17	41-44	42	61-68	174-185	189-198	341-390
16	39-40	40-41	55-60	162-173	175-188	315-340
15	34-38	35-39	49-54	147-161	162-174	299-314
14	30-33	32-34	42-48	134-146	142-161	256-298
13	24-29	27-31	40-41	114-133	123-141	233-255
12	17-23	23-26	33-39	89-113	105-122	190-232
11	13-16	19-22	27-32	75-88	89-104	158-189
10	9-12	15-18	22-26	63-74	71-88	136-157
9	7-8	12-14	16-21	54-62	58-70	115-135
8	5-6	8-11	9-15	45-53	36-57	93-114
7	-	6-7	5-8	34-44	26-35	64-92
6	3-4 (4)	5	2-4	27-33	19-25	49-63
5	1-2	3-4	1	18-26	15-18	36-48
4	0	2	0	13-17	7-14	21-35
3	0	1	0	8-12	4-6	11-20
2	0	1	0	3-7	3	6-10
1	0	-	0	1-2	1-2	1-5
0	0	0	0	0	0	0

Table 2

254 SAMPLE, ITEM STATISTICS

Subtest	Information		Comprehension		Arithmetic		Analogies		Vocabulary	
mean	10.551		14.558		6.870		7.724		29.712	
S.D.	7.286		6.421		3.128		5.950		19.791	
Item No.	f of non-zero answers	Item Test r	f	r	f	r	f	r	f	r
1	244	.231	243	.063	252	.139	176	.396	205	.480
2	246	.182	231	.407	252	.153	219	.401	207	.465
3	239	.247	228	.268	234	.322	165	.417	157	.553
4	200	.389	178	.489	203	.388	127	.557	204	.529
5	135	.488	165	.374	192	.446	140	.580	195	.561
6	157	.447	229	.434	120	.471	149	.409	144	.615
7	110	.609	134	.506	104	.481	63	.602	194	.654
8	131	.544	123	.563	75	.556	80	.579	131	.610
9	70	.583	128	.629	80	.472	67	.778	196	.543
10	112	.494	176	.514	65	.592	40	.627	140	.615
11	80	.653	146	.484	18	.389	58	.591	170	.540
12	113	.631	143	.535	49	.514	15	.390	185	.491
13	94	.562	209	.574	21	.421	37	.534	155	.534
14	54	.686	99	.557	30	.536	26	.595	142	.497
15	101	.655	69	.481	20	.424	23	.451	76	.643
16	47	.713	66	.547	13	.487	10	.334	145	.559
17	71	.711	131	.475	6	.276			121	.096
18	22	.585	16	.440					128	.589
19	19	.521							149	.566
20	60	.703							156	.509
21	41	.650							166	.558
22	10	.463							160	.418
23	60	.705							162	.532
24	37	.539							148	.618
25	28	.596							56	.628
26	19	.582							60	.532
27	19	.566							64	.718
28	19	.458							127	.452
29	48	.603							68	.663
30	37	.598							81	.642
31	19	.570							75	.522
32	15	.490							98	.591
33	7	.380							57	.631
34	12	.420							92	.658
35	9	.402							100	.582
36	4	.217							71	.634
	There was only one zero score		There was only one score of zero		There were some failures to actually administer items 1 and 2 when needed		There were seven scores of zero			

Table 2

(Cont.)

Subtest:			Picture Comp.		Block Design		Picture Arr.		Obj. Assembly	
Mean	Vocabulary		10.535		15.488		14.338		17.856	
S.D.	(Cont.)		6.422		10.148		10.504		10.34	
Item No.	f	r	f	r	f	r	f	r	f	r
1			185	.460	214	.487	238	.334	254	.521
2			210	.348	165	.590	234	.382	223	.542
3			142	.519	174	.516	106	.569	192	.580
4			97	.472	168	.532	43	.545	135	.566
5			143	.514	121	.677	40	.545	158	.451
6			90	.324	80	.611	53	.628		
7			90	.620	63	.635	67	.638	There were no zero scores	
8			108	.507	18	.396	17	.302		
9			49	.542	19	.411	43	.489		
10			139	.623	6	.270	18	.347		
11			173	.582			78	.689		
12			44	.504	There were twenty scores of zero		41	.365		
13			9	.329						
14			42	.412			There were eight scores of zero			
15			159	.420						
16			46	.544						
17			57	.524						
18			155	.570						
19			28	.322						
20			13	.354						
21			17	.313						
22			96	.507						
23			87	.569						
24			81	.552						
25			47	.302						
26			108	.576						
27			167	.371						
28			115	.633						
37	46	.533	There were eight scores of zero							
38	62	.688								
39	60	.494								
40	45	.678								
41	39	.559								
42	58	.602								
43	14	.377								
44	31	.613								
45	23	.496								
46	32	.554								
47	15	.444								
48	9	.391								
49	15	.477								
50	3	.298								
51	4	.298								
	There were seven scores of zero									

Digit Span	Digit Symbol
mean & S.D.	mean & S.D.
Approx. 7.6 Approx. 2.21	Approx. 24.9 Approx. 16.3

All this led to an investigation of various subdivisions of the data. After all, the English language WAIS was standardized on a population in which illiteracy was lower and average education was higher. Also the scaled scores were set up on the age bracket which earned the highest scores. This suggests (a) that the English WAIS might not scale too well if the total population were used and, therefore, (b) that the Spanish WAIS might scale much better or even satisfactorily if a restricted portion of the population were used to establish such scores. As long as the reference group is well defined this might be a reasonable approach. As an aside, however, it should be mentioned that the use of the highest scoring age bracket to set up scaled scores is likely to have some effects that tend to be underemphasized in the States. Perhaps such effects are not too important but the present writers do not recall seeing any discussion of them anywhere.

2. Subdivisions of Education

Suppose, for instance, we separate out all subjects who had three years of education or less. It is by no means certain that the WAIS will work well with illiterates. At least experience with such has not been great in the States.

In Table 3 a summary of such a division is presented. The means, standard deviations and highest scores of the 192 subjects who had had four years of education or more are compared with the same statistics of the 62 subjects who have had three years of education or less. The means and standard deviations of the total sample are also included in the table for convenience. Only the nine subtests having items in the usual sense are included.

It is clear that this subdivision makes a large difference in mean performances on all nine subtests included in the table. Also, except for Item 1 of Object Assembly all standard deviations in the low education group are smaller. Generally, the degree of skewness of the various subtests appears less in the more educated group. But only for Arithmetic and Comprehension does the distribution approach symmetry. In all other subtests skewness is clearly present, the means typically being less than two standard deviations above zero. The fact that these tests typically do not scale well was verified by setting up scaled scores in the manner of Table 1.

An important consideration here concerns the question as to whether or not item-test correlations are satisfactory for the low education group. The right hand portion of the lower part of the Table 3 refers to this question. It will suffice here to say that for almost all items on which even two or three people earn a score,

Table 3

192 SUBJECTS WHO HAVE HAD 4 YRS. OF ED. OR MORE

Compared with total sample of 254 and with 62 subjects who have had 3 years of Ed. or less

192	M	S.D.	Highest Score	254	M	S.D.
Vocab.	33.859	20.021	92 of 102		29.712	19.791
Arith.	7.343	3.096	17 of 17		6.870	3.128
Infor.	12.083	7.509	32 of 36		10.551	7.286
Analogies	8.848	6.210	30 of 32		7.724	5.950
Comp.	15.828	6.320	33 of 36		14.558	6.421
P.C.	12.171	6.088	27 of 28		10.535	6.422
B.D.	17.901	9.545	44 of 48		15.488	10.148
P.A.	16.515	10.762	46 of 60		14.338	10.504
O.A.	19.993	10.11	46 of 55		17.856	10.34

SUBJECTS WHO HAVE HAD 3 YRS. OF ED. OR LESS N = 62

			Highest Score	No. of Items which Correlate with Subtest	Last Item which Correlate with Subtest
Vocab.	16.87	12.048	42 (next highest)	28	34
Arith.	5.403	2.750	15 (10)	8	10
Infor.	5.806	3.618	17	12	15
Analogies	4.241	3.093	15	8	9
Comp.	10.612	4.968	20	13	17
P.C.	5.661	4.885	21 (16)	3	25
B.D.	8.016	8.121	29 (24)	6	6
P.A.	7.596	5.762	34 (21)	4	7
O.A.	11.352	7.99	38 (27)		

the correlations tend to be .3 to .6 and sometimes even higher. So the answer clearly is yes; for people of very low education, at least in an area in which low education is in many cases likely to reflect little opportunity, a sufficient number of the test items function normally.

3. Subdivisions of Age

The next attempt was to investigate roughly what would happen if subjects were subdivided according to age. Since it can be presumed that optimum scaling was achieved for the English WAIS by using the age bracket they did (20-34), perhaps a reasonable approximation to the same effect would occur here also. Since, however, the number of years of education is increasing rapidly, this would mean that our optimal age bracket is likely to be lower than that for the English WAIS. We did not have enough subjects to investigate what the optimum bracket might be and so elected to divide the subjects into three subdivisions: 16-24, 25-44, and 45-64. Comparisons of the means, standard deviations and the frequency of zero scores of nine subtests are presented in Table 4 for these three age divisions.

Again skewness in most subtests is evident. In no case would scaling be possible for more than three of the nine subtests but it clearly would be better overall at one of the two younger age groupings.

There is, however, a result illustrated in Table 4 which is important to note in evaluating the meaningfulness of the project. The average performances are in accordance with outcomes observed in the States. In the verbal subset of the subtests, the middle group scores highest on every subtest (although just barely on Arithmetic), the oldest group comes in second in every case, and the youngest group comes in last on all five of these subtests. On all four performance subtests the youngest group is highest in every case, the middle group is next, and the oldest group is lowest on all. The standard deviations do not show these trends so consistently although they do appear to be somewhat correlated with the means.

It was clear at this point that no age division would be likely to produce satisfactorily scalable subtests although it would increase somewhat the number of subtests that would be approximately scalable. It remained possible, of course, that restricting both age and education would produce a scalable group. Since the test functions properly for low education groups, it would hardly be appropriate to restrict the norms to some minimum standard of education. To do so would be to give an arbitrary decrease in average ratings to low education people. It would seem that some other scheme should be devised if one wishes in some way to make allowances for lack of opportunity.

4. Other statistical characteristics of the test based on the 254 sample.

Even though we were off the mark on difficulty, it seemed to be essential to the question of meaningfulness of the project that we go ahead investigations of other properties of the test as it found that some of these might compare unfavorably with similar indices on the English WAIS. But also, since skewness was in the same direction for all distributions, such indices should not prove too badly affected. So, if other indices had been seriously lower than corresponding ones for the English WAIS, then we would have had to consider that the idea of applying the same definitions of intelligence in the two cultures was probably invalid. Actually, virtually all comparisons we have made to date are quite favorable. They will be reviewed in turn.

a. Reliabilities of the subtests

Split half reliabilities, corrected for length, have been computed on nine of the subtests. The items were ordered according to difficulty. An estimate similar to that used for the English WAIS was used for Digit Span--digits forward was correlated with digits backwards and then corrected for length. These reliabilities, except possibly for digit span, are surprisingly good. The estimates for digit span average somewhat lower than for the English version. All the others are either as high or somewhat higher than for the English WAIS. They are presented in Table 5.

No reason for the exception noted for Digit Span is immediately apparent. Some zero scores did occur which produced a discontinuity in the scatter plots between 0 and 3 for digits forward and between 0 and 2 for digits backwards. Such a discontinuity should be no more likely to attenuate correlation than it is to spuriously inflate it.

b. Intercorrelations of the 11 subtests among themselves and with Verbal, Performance, and Full Scale raw scores.

The intercorrelations among the 11 subtests are presented in Table 6. They have been calculated only for the total sample.

Table 4

COMPARISON OF MEANS, STANDARD DEVIATIONS, AND FREQUENCY OF SCORES OF 9 SUBTESTS OF 3 AGE DIVISIONS

Age	16-24			25-44			45-64		
N	83			96			75		
Statistic	mean	S.D.	O'S	mean	S.D.	O'S	mean	S.D.	O'S
Subtest									
Information	9.20	6.23	(3.1's) 0	11.83	8.03	(4.1's) 0	10.40	7.08	1
Comprehension	13.29	5.81	0	15.83	6.26	0	14.32	6.94	1
Arithmetic	6.30	2.66	(5.2's) 0	7.18	3.03	0	7.11	3.60	1
Analogies	6.52	5.10	(6.1's) 2	8.72	6.04	(5.1's) 1	7.79	6.44	4
Vocabulary	25.83	18.04	2	33.40	19.34	1	29.29	21.40	4
Picture Comp.	12.34	6.25	0	11.31	5.95	3	7.65	6.15	5
Block Design	17.66	10.10	4	16.00	8.95	5	12.43	10.89	11
Picture Arr.	17.12	10.40	3	15.01	10.36	2	10.40	9.13	3
Object Assembly	21.46	8.43	0	17.67	13.54	0	14.07	11.72	(3.1's) 0

Table 5

RELIABILITIES OF TEN SUBTESTS FOR ALL SUBJECTS AND FOR THREE AGE GROUPINGS

Subtest	Total 254	16-24 83	25-44 96	45-64 75
Information	.85	.90	.77	.92
Comprehension	.86	.80	.85	.90
Arithmetic (no time bonus)	.88	.90	.85	.91
Analogies (Simil.)	.88	.87	.87	.91
Digit Span	.57	.66	.57	.55
Vocabulary	.95	.96	.95	.93
Picture Completion	.90	.89	.86	.91
Block Design	.87	.88	.83	.90
Picture Arrangement	.86	.87	.84	.84
Object Assembly (5 items)	.77	.78	.68	.78
Object Assembly (4 items)	.75			

Table 6

INTERCORRELATIONS OF THE ELEVEN SUBTESTS

Total Sample - 254
(Using Alternate Scoring)

	Information	Comprehension	Arithmetic	Similarities	Digit Span	Vocabulary	Digit Symbol	Picture Completion	Block Design	Picture Arrangement
Comprehension	.805									
Arithmetic	.684	.642								
Similarities	.824	.751	.683							
Digit Span	.553	.544	.592	.527						
Vocabulary	.853	.797	.653	.802	.584					
Digit Symbol	.593	.574	.529	.543	.576	.570				
Picture Completion	.634	.602	.569	.571	.563	.609	.726			
Block Design	.623	.626	.630	.564	.599	.598	.732	.736		
Picture Arrangement	.679	.608	.525	.570	.503	.640	.728	.781	.709	
Object Assembly	.558	.496	.475	.483	.509	.521	.640	.682	.701	.671

Table 7

INTERCORRELATIONS OF THE 11 SUBTESTS WITH VERBAL, TOTAL,AND FULL SCALE
RAW SCORE TOTALS, AND ALSO OF VERBAL AND PERFORMANCE WITH FULL SCALE N = 254*

	Information	Comprehension	Arithmetic	Analogies	Digit Span	Vocabulary
Verbal	.893	.836	.715	.847	.614	.882
Performance	.701	.660	.617	.620	.627	.667
Full Scale	.840	.790	.712	.770	.670	.780

	Digit Symbol	Picture Comp.	Block Design	Picture Arrangement	Object Assembly	Verbal Score	Performance Score
Verbal	.626	.664	.666	.681	.567		
Performance	.796	.831	.814	.812	.744		
Full Scale	.756	.807	.797	.805	.704	.917	.940

* All correlations have been corrected for over-lap.

Again one sees that the correlations tend to be as high and usually higher than those obtained for the English WAIS. The tendency for these to be higher could be due to the use of the total sample whereas for the English WAIS these correlations are presented for age restricted groups.

Intercorrelations of the 11 subtests with Verbal, Performance, and Full Scale Raw scores are presented in Table 7. And still one more time they are seen to be comparable to those reported in the manual for the English (U. S.) WAIS. The greater part of them are slightly higher and only a few are slightly lower. However, it should be noted that all coefficients in Table 7 are based on the total sample whereas those reported in the WAIS manual are based on those subidivisions of ages. It is not clear, therefore, whether or not the actual correlations reported here are in fact higher in general than their counterparts for the U. S. WAIS.

c. Validity of the Test

Much of the evidence presented so far can be considered indirect evidence for the validity of the test. If its properties are the same in the two cultures, then it may well prove to have predictive efficiency for the same criterion variables.

Of course, at this time there was no direct validity evidence of any kind for the test that was being formed. There is another indirect bit of evidence that suggests that validity for predicting success in school may prove to be fairly good. Correlations between number of years reported by the subjects on the one hand, and each of the fourteen subtest and total scores obtained from the WAIS on the other have been calculated. These, are presented in Table 8. Since these correlations are as high as comparable coefficients in the States it may well be that school success will be predicted as well here as there.

Actually there are logical grounds on which to predict that less validity will be achieved. These stem from cultural differences in attitude toward school achievement that appear to be real and significant. These cultural differences do not preclude validity. After all, all are in the same culture and so this variable may effectively be cancelled. It will be interesting some day to find out what the answer really is.

If the validity for predicting school success achieved is the same as it is in the States it may be for somewhat different reasons. Performance scores are usually thought of as being somewhat more independent of formal education than verbal scores. This does

Table 8

CORRELATIONS* BETWEEN NUMBER OF YEARS OF EDUCATION REPORTED AND WAIS RAW SCORES

VERBAL		PERFORMANCE	
Information	.60	Digit Symbol	.70
Comprehension	.42	Picture Completion	.51
Arithmetic	.33	Block Design	.48
Analogies	.42	Picture Arrangements	.57
Digit Span	.46	Object Assembly	.42
Vocabulary	.42		
Ave. of the 6 Phi's	.44	Ave. of the 5 Phi's	.54
Total Verbal	.68	Total Performance	.68
Full Scale	.71		

* Phi Coefficients for the 11 sub-tests, education variable split between 6th and 7th grades or a 48% versus 52% split.

* Pearson r's for the three total scores

not happen in the Puerto Rican population. In fact there is an opposite tendency involved. The performance subtests, as a set, tend to correlate with education higher than the verbal subtests. The lowest correlation is .33 with Arithmetic, a so-called verbal test, and the highest is .70 with Digit Symbol Substitution, a so-called performance test. If this pattern of correlations is duplicated in the standardization sample, then this may well prove to be one of the few areas of difference between populations that will be found.

Finally, one should not fail to note that even for the highest correlations obtained, .70 on Digit Symbol and .71 on Full Scale Total, only 50% of the test variance is associated with formal education. Obviously, although all subtests correlate significantly with education, there are important amounts of variance that are not directly associated with formal education. It would be of interest to determine what, if anything, this variance is associated with. For instance, if the variance due to education were partialled out, would the resulting scores be better predictors of probably benefit from special training programs? Might this not be particularly true in a population in which opportunity for formal education is still somewhat limited? An attempt will be made to investigate this possibility.

5. Special analyses other than difficulty that have been investigated with the 254 sample data.

a. Sex differences

It will be recalled that there were statistically significant sex differences observed on 8 of the 11 subtests of the English WAIS. These differences were calculated on the total standardization sample. A question that arises concerns whether or not similar differences exist in the Puerto Rican culture. If they are in fact observed, this too might lend some indirect evidence for the validity of the Spanish WAIS.

The sex differences obtained in raw scores are presented in Table 9. It will be noted that on average, males scored higher than females on all three of the total scores and also on all eleven of the subtest scores. Differences occurred on eight subtests in the English WAIS with five favoring males and three favoring females, and therefore virtually vanished in the Full Scale total.

An immediate hypothesis concerning the source of this difference is that Puerto Rican males may get more schooling than females. And it turned out that the average male in the sample had reported having completed 7.7 years of school while the comparable figure for females was 6.7. Since the correlation between education and total score is about .70, the sex difference in scoring might be largely if not totally explained by the educational difference.

The subsequent sample of 124 subjects (which is described in detail a little later) used for the later effort to inject some easier items into the test proved to have the educational status of the sexes almost exactly equal. The men reported 7.81 and the women reported 7.75 years of education. This shift probably was due to our having taken the 124 sample from census district #1 which is urban and rural coastal. It is likely that the educational differences occur largely in the "mountainous" interior of the island.

At any rate, the sex differences in the 124 sample for 8 subtests are presented in Table 10. The differences reported are not of themselves statistically reliable. Nevertheless, in absolute values males exceed females on 6 and females exceed males on 2. One of these two is Vocabulary. This fact is quite comforting since the sex difference in verbal development has been long known and widely observed. It certainly lends credence to the growing belief that this test will in fact function at least adequately in this culture. The other subtest on which women excelled men is Digit Symbol Substitution. This too fits our expectations from experience with the English language WAIS. It occurred in spite of the fact that 7 women earned scores in the 0-4 range while only 3 men did so.

b. A brief look at Urban-Rural differences

Urban-rural differences have been investigated in three ways. First the total sample was divided into the two categories; then persons with 0, 1, or 2 years of education were omitted; finally two subsamples were compared in which number of years of school and the examiner were controlled.

The data relevant to the first two of these is presented in Table 11. Subtantial differences between averages of Urban and Rural examinees occurred. One would have expected these differences to be more marked in the total samples than in the subsamples having 3 or more years of education. However, dropping the 30 rurals with 0, 1, or 2 years of education versus dropping the 12 similar urbans leaves differences between means largely unchanged and only slightly less in any case.

Two further observations from Table 11 should be highlighted. First, note that most of the means of the score distributions of the urbans with three years or more of education are two or more

Table 9

SEX DIFFERENCES IN TOTAL SCORES

N = 254

	Males		Females	
	M	S.D.	M	S.D.
Verbal	82.02	40.50	70.44	37.12
Performance	92.23	48.33	81.59	45.92
Total	172.37	82.39	151.99	77.60

Verbal Subtests

	Males		Females	
	M	S.D.	M	S.D.
Information	12.54	8.30	9.10	5.88
Comprehension	15.87	6.43	13.32	5.50
Arithmetic	7.64	3.12	6.14	2.94
Analogies	8.79	6.30	6.79	9.61
Digit Span	7.91	1.96	7.41	2.21
Vocabulary	31.56	16.00	28.44	14.13

Performance Subtests

	Males		Females	
	M	S.D.	M	S.D.
Digit Symbol	25.84	16.34	23.96	15.79
Picture Compl.	11.63	6.35	10.32	5.98
Block Design	17.49	9.70	13.79	9.48
Picture Arragt.	18.18	12.02	16.15	11.62
Object Assembly	19.19	8.75	16.49	9.91

standard deviations from zero and hence are not too far from being scalable. Second, even for this group, there are only three subtests on which the means are essentially at the middle of the theoretical range of scores. These are Arithmetic, Comprehension, and Picture Completion. Arithmetic, however, has no time bonus involved in its scoring at this point.

The third look at Urban-Rural difference stemmed from the possibility that simple differences in number of years in school or that the sex of the tester might be influencing the results. The sample sizes were small and the investigation was done only on raw score total. Even so, for the two examiners a total score difference of 35 points for the male and 20 points for the female favoring urban over rural was found. This difference probably reflects differential quality of schools in the two zones.

B. The Guatemala Extension of the Project

While the testing of the 254 Sample was still in progress, members of the Department of Psychology of the University of Guatemala asked that they be permitted to test a sample of people in that Country. This was deemed a desirable extension of the project and an application for supplemental funds was approved by the Office of Education.

The Guatemala extension was limited to Guatemala City, within which probably 80% of the spanish speaking people of the country live. The sampling methods used were the same except that the samples were drawn though the cooperation of the Bureau of the Census. The total sample tested consisted of 293 subjects. These subjects were all urban dwellers, they had an average education of almost exactly 5 years, they averaged 33.0 years of age, and 48.2% of them were male. The urban portion of the Puerto Rican sample, by contrast, had an average of 8.5 years of education, averaged 35 years of age, and also had some 48% males. The higher average age reflects the relatively higher longevity in Puerto Rico, and the higher average number of years of school reflects the higher literacy rate achieved in Puerto Rico as of now.

The Guatemala data was subjected to virtually every analysis carried out on the Puerto Rican 254 (item-analysis) sample. The over-all item-analysis, subtest intercorrelation, and reliability results were so similar that to present detailed tables would be redundant. Probably the most noteworthy difference observed was that even through they had had an average of 3.5 years less of education, the Guatemalan sample scored as well as the Puerto Rico

Table 10

SEX DIFFERENCES ON 8 SUBTESTS 124 SAMPLE

Verbal Subtests

	Males (n = 53)		Females (n = 71)	
	Mean	S.D.	Mean	S.D.
Information	19.30	7.10	17.66	7.17
Arithmetic	9.79	3.19	8.15	2.89
Analogies	13.26	6.65	12.61	6.63
Vocabulary	41.38	18.95	42.70	19.38
Sum	83.73		81.12	

Performance Subtests

Digit Symbol	32.74	16.95	35.57	21.15
Picture Completion	15.87	5.56	14.96	5.71
Block Design	20.49	10.74	18.39	11.48
Picture Arrangement	28.89	15.99	27.44	15.56
Sum	97.99		96.36	

Table 11

URBAN-RURAL DIFFERENCES

	Urban Ed: all		Rural Ed: all		Urban Ed: 3+		Rural Ed: 3+	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Vocabulary	37.17	19.96	23.64	17.46	39.05	19.98	26.74	17.59
Arithmetic	7.69	3.49	6.20	2.61	7.96	3.46	6.57	2.45
Information	13.60	7.74	8.07	5.84	14.39	7.73	9.07	6.09
Analogies	9.72	6.62	6.10	4.77	10.36	6.64	6.69	4.94
Comprehension	17.37	6.67	12.26	5.58	17.93	6.13	13.19	5.57
Pict. Comp.	12.64	6.60	8.82	5.73	13.70	6.29	9.98	5.46
Block Design	18.61	10.02	12.94	9.52	20.08	9.40	14.96	9.23
Pict. Arr.	18.56	11.69	10.90	7.91	19.88	11.53	12.16	8.31
Obj. Assembly	20.54	18.58	15.68	9.57	21.67	10.34	17.57	9.60

Full Scale Differences with groups matched on education
and with Sex of Tester held constant.

	Male Tester		Female Tester	
	N	Mean	N	Mean
Urban	21	164	13	151
Rural	26	129	13	131

urban sample on Vocabulary, Arithmetic, Block Design, Picture Arrangement, and Object Assembly. They scored lower than the Puerto Rican sample on Information, Comprehension, Analogies, and Picture Completion. When Guatemalan subjects having 4 years or more of education are compared with Puerto Rican subjects also having 4 or more years (but both Urban and Rural subjects being included) then the Guatemalan sample scored higher on all subtests except Comprehension. In fact the test difficulty level was almost appropriate for the Guatemalan subjects of 4 years and above in education.

There were, of course, some differences in item difficulty levels, especially in the vocabulary subtest. Even so, in so far as item, test correlation were concerned there was virtually complete over-lap in the sense that with only one or two exceptions, items that were satisfactory in Puerto Rico were also satisfactory in Guatemala.

The Guatemalan males scored higher, on average, on every subtest than the females. This differences probably is primarily due to differences in educational levels achieve by the sexes.

The age pattern of scores was very similar to that in Puerto Rico. The middle (25-39) age group was somewhat higher on 5 verbal tests than the oldest (40-64) group which in term was about equal to the youngest (16-24) group. Digit span decreased from youngest to oldest as did every performance sub-test.

C. Development of new manual and new Item Analysis, sample no. 2
N = 124, all from Census Region #1

After consultation with the Psychological Corporation, it was decided to try out a series of new and easy items in order to try to achieve better discrimination among low scoring people on eight of the subtests. Comprehension was judged reasonably acceptable as it was, Object Assembly was perhaps not quite as good but considered relatively acceptable, and it was considered that Digit Span could not reasonably be modified.

We, therefore, constructed new items for all other subtests. A brief review of some of the considerations that guided this effort would probably help here. Our item difficulty indices showed that we had very few items in any subtests on which as many as 80 to 90 per cent of the population could succeed. Our efforts were in all cases to construct new material which might place items in the 85 to 95% success range.

A constant criterion that was kept before the group was that we should search very hard for items which would be concerned with things that should be universally experienced. It was this criterion that led to such questions in Information as number 5, "why do birds have wings?" and number 9, "what color is grass?". On number 5, we obtained 100% correct answers and on number 9, we obtained 96.8% correct answers. These two plus the original item, "what is the shape of a ball?", now form the first three items of the standardization form with automatic credit for anyone who gets 4 and 5 correct. Item 4 (8 for the 124 sample) is "from what plant is sugar obtained?" and was answered correctly by 96% while item 5 (no. 1 in 124 sample) has a difficulty index of about 94%. This criterion, then, seems to have been sufficient to get us through the information test reasonably well.

A word of caution needs to be inserted here, however. When the decision to try again was made it was also decided to do it on a smaller sample and to restrict the sampling to Census Region #1. Census Region #1 contains San Juan and is otherwise more coastal than mountainous. This sample was, therefore, somewhat biased in the direction of being slightly better able to respond to this test than would have been true of the whole island. This bias in the sample plus the fact that the sample was relatively small probably meant that these items would be expected to be found to be a little more difficult in the total population than these figures indicate.

The Vocabulary subtest also merits some special discussion. For item selection we used as our main guides, universality and frequency of use. We included new words which were almost exclusively among the 100 most used words in the language. This was our third try at finding words which would permit us to give automatic credit. In the two previous tries we never obtained scorable responses from more than about 80% of the subjects on any word. This time we reached 90% with gato (cat), 91% with finalizar (to finish, end) and 92% with vaso (glass, usually to drink from). We do not, of course, believe that these simple words are not known in the sense that they are not a part of the meaningful and used vocabularies of the low-scoring people. It is that they just cannot respond to the question of what the words mean. For example, we are assured at times that "a gato es un gato" (a cat is a cat). We score it zero but the person almost surely knows what a cat is, i.e., knows the referent for the symbol. Again, the per cents reported above (except for vaso, which was not used for the standardization form) were expected to drop some when sampling was island-wide again. All examinees are started with item number one in the vocabulary subtest.

With this much background, then, let us review some of the highlights of the data analysis of this sample.

1. Scaling, Difficulty Levels, and Item Subtest correlations

Since the sample was so small, scaling at the ends of the distribution could not be done satisfactorily. However, for comparative purposes a scaling was done and the results are presented in Table 12. Clearly, five of the sub-tests did scale better at the the low end of the scale. The three poorest tests were Digit-Symbol, Analogies and Block Design. Others, while improved, still did not scale as well among low scores as would have been desirable. We felt that the test had been overloaded with easy items. (In fact test administrators often complained that the test was too easy.) In spite of this, Picture Completion was the only subtest that was in fact somewhat too easy (for the subset of the population from which we sampled, that is).

Item difficulties and Item-Subtest correlations (corrected for overlap) are presented in Table 13.

For Information and Picture Completion we had a reasonably good distribution of difficulty. Other subtests had one or more questionable features. Picture Arrangement had too few items in the fairly easy range. We had inserted a series of (we thought) very easy 4-card items. The difficulty index results were rather disappointing although the item test correlations were all fine. To make the difficulty index of this subtest any better apparently would have required adding at least two more 3-card items, if in fact that would have done it. Twenty-four people did not earn credit on the first attempt on item 1 and 23 people did not do so on item 2. Perhaps continuing to give two trials for more items would have also been required.

Block Design also had an empty section in the difficulty index distribution. Here we felt we had probably used up variants of the easy patterns that would not be near duplications. It seemed highly unlikely that the empty section could be filled.

Vocabulary was good except for our failure to get above 91 or 92 per cent. As of now we are forced to admit that there seems to be doubt that it can be done.

Analogies looked fairly good except that in this case, even though the difficulty distribution looked pretty good, we got four scores, or three-plus per cent, of zero. We are at a loss as to how this could be overcome.

Table 12

SCALED SCORES FOR THE SECOND SAMPLE OF 254 SUBJECTS

Scaled Score	Prop.	Arithmetic	Information	Analogies	Vocabulary	Digit Symbol	Picture Completion	Block Design	Picture Arrangement	These scales are sums of raw scores		
										V	P	FS
19	.999		37	29-34	86-100	82-110	25-27	46-48	60-68	162-188	197-253	342-441
18	.998	17	35-36	(28)	(85)	(81)		(45)	59	(154-161)	(194-197)	(338-341)
17	.994		34	27	84	(79-80)	24	44	56-58	153	190-193	330-337
16	.985	15-16	33	26	81-83	71-78		42-43	54-55	149-152	184-189	317-329
15	.967	14	31-32	24-25	77-80	64-70		39-41	52-53	145-148	159-183	304-316
14	.933	13	29-30	21-23	67-76	57-63	23	33-38	50-51	114-144	154-158	289-303
13	.878	12	25-28	19-20	60-66	52-56	21-22	29-32	42-49	113-125	143-153	251-288
12	.798	11	22-24	16-18	51-59	44-51	19-20	25-28	39-41	99-112	126-142	221-250
11	.691	10	19-21	14-15	43-50	48-43	18	21-24	33-38	81-98	105-125	177-220
10	.566	8-9	17-18	12-13	36-42	30-37	16-17	18-20	25-32	70-80	86-104	158-176
9	.434	7	14-15	9-11	31-35	23-29	13-15	14-16	17-24	61-69	72-85	140-157
8	.309		12-13	7-8	27-30	17-22	11-12	10-12	13-16	57-60	55-71	110-139
7	.202	6	11	6	23-26	12-16	8-10	6-8	10-12	46-56	39-54	87-110
6	.122	5	9-10	5	16-22	3-11	6-7	4	6-8	37-45	24-38	69-86
5	.067		8	1-4	9-15	1-2	5		4	27-36	16-23	62-68
4	.033	4	6-7	0	6-8	0	4	2	2	(19)-26	13-15	15-61
3	.015	3	5		1-5		3	0		7-(18)E	7-12	13-54
2	.006	0	4		0		(2)E		0	6	6	12
1	.002	?	(2-3)Est		?		(1)E		?	?	?	?
0	.001	0	(0-1)Est	0	0	0	(0)E	0	0	?	?	?

Table 13

124 SAMPLE, ITEM STATISTICS

Subtest:	Information:		Arithmetic:		Analogies:		Vocabulary:		Picture Completion:		Block:		Picture Arrangement	
mean	18.26		8.85		12.89		42.14		15.35		Design		28.06	
S.D.	7.24		3.13		6.64		19.21		5.66		19.29		15.76	
Item											11.21			
No.	D	r	D	r	D	r	D	r	D	r	D	r	D	r
1	117	.304	123	.228	97	.362	111	.345	117	.254	114	.387	122	.442
2	121	.167	122	.262	90	.492	81	.514	116	.397	111	.442	116	.546
3	96	.208	115	.341	84	.636	112	.347	118	.332	78	.508	55	.494
4	86	.248	117	.301	116	.299	110	.443	117	.413	52	.570	76	.531
5	124	.000	117	.197	105	.350	105	.471	111	.413	87	.605	77	.646
6	116	.100	109	.274	108	.375	114	.343	95	.392	61	.641	40	.546
7	27	.329	100	.438	111	.444	91	.334	102	.540	47	.666	66	.596
8	119	.203	53	.606	86	.604	39	.414	101	.432	33	.609	55	.431
9	121	.203	59	.608	71	.467	106	.611	85	.403	12	.489	79	.659
10	90	.270	32	.538	80	.617	113	.409	111	.230	6	.380	36	.569
11	97	.215	39	.614	70	.671	111	.235	53	.461			45	.484
12	104	.448	37	.572	55	.547	95	.582	39	.479			35	.435
13	88	.403	34	.634	30	.435	111	.441	50	.559			32	.511
14	81	.425	8	.398	44	.712	76	.533	73	.471			32	.408
15	67	.612	12	.436	36	.632	87	.608	34	.480	There			
16	71	.568	12	.427	22	.496	97	.547	71	.664	were			
17	56	.562	9	.487	19	.513	105	.613	41	.403	three		There was	
18	78	.459					92	.517	55	.627	scores		one score	
19	56	.669					79	.668	47	.493	of zero		of zero	
20	97	.441	there was		There were		114	.548	74	.489				
21	51	.625	one score		four scores		100	.569	41	.465				
22	51	.606	of zero		of zero		93	.462	20	.378				
23	32	.660	and two				50	.584	45	.348				
24	39	.642	scores				101	.607	61	.520				
25	34	.579	of 17				53	.696	28	.454				
26	26	.589					91	.511	84	.616				
27	30	.643					103	.571	14	.116				
28	36	.575					90	.524						
29	23	.608					86	.501	There were					
30	18	.427					85	.455	no scores					
31	35	.577					28	.616	of zero					
32	16	.394					46	.493						
33	9	.398					58	.552						
34	20	.616					96	.534						
35	14	.594					42	.587						
36	8	.505					88	.563						
37	9	.433					47	.567						
38							44	.622						
39							62	.367						
40	There were						43	.725						
41	no scores						39	.618						
42	of zero						62	.513						
43	and one						51	.686						
44	score of						51	.523						
45							41	.571						
46							24	.582						
47							36	.596						
48							32	.659						
49							20	.552						
50							25	.526	There was one score of zero					

Digit Symbol

Mean 34.29
S.D. 19.28

Range 0-84

	Verbal Total	Performance Total	Full Scale
X	80.89	97.5	178.92
S.D.	36.72	34.79	75.35
Range	6-162	6-197	12-342

Arithmetic also suffered from the "empty middle syndrome" in difficulty indices. The instant one introduces problems that require anything more than simple addition or subtraction the index takes a deep dive. It just might be that if we had inserted items involving a series of more than three additions and subtractions we might have been able to get something in that space. It would be interesting to try sometime. However, we were boxed in on this subtest unless we were to lengthen it still more. The U. S. WAIS used 14 items. This form used 17. Apparently an optimum length here would have been more like 20.

Finally, Digit-Symbol looked essentially hopeless. It is clear that simplifying this test did not help much.

2. Inter Subtest Correlations and Correlations of Subtests with Verbal, Performance, and Full Scale raw score sums

The subtest intercorrelations are presented in Table 14. They are quite consistent with those from the 254 sample, or if anything they average a little higher.

The correlations of the subtests with Verbal, Performance, and Full Scale raw score sums, corrected for overlap, are presented in Table 15. These too are quite consistent with those from the 254 sample. It is also apparent that the Verbal subtests correlates higher with the Verbal total than the Performance subtests do, and vice versa.

It was quite clear that in these respects all of these--the subtests and total scores--were functioning according to expectations.

3. Subtest score distributions

One might wonder what effects the kind of difficulty distributions reported for Picture Arrangement, Block Design, and Arithmetic might have had on score distributions. There were no easily demonstrable effects as Table 16 will illustrate. Distributions were fairly regular although with some tendency to skewing in some cases. Vocabulary probably was rather platykurtic but, on the whole, the score distributions looked better than had been expected.

D. The decision to go ahead to the standardization sample and the item selection process.

At this point, we were faced with the decision as to what to do

Table 14

SUBTEST INTERCORRELATIONS - 124 SAMPLE

	Information	Arithmetic	Analogies	Vocabulary	Digit Symbol	Picture Completion	Block Design
Arithmetic	.678						
Analogies	.803	.676					
Vocabulary	.869	.684	.811				
Digit Symbol	.683	.571	.628	.684			
Picture Completion	.729	.510	.604	.630	.763		
Block Design	.675	.533	.636	.627	.730	.709	
Picture Arrangement	.614	.462	.609	.596	.753	.768	.726

Table 15

SUBTEST CORRELATIONS WITH V, P, AND FS, CORRECTED FOR OVERLAP

124 SAMPLE

	Information	Arithmetic	Analogies	Vocabulary	Digit Symbol	Picture Completion	Block Design	Picture Arrangement	V	P
Verbal	.853	.720	.828	.796	.697	.660	.670	.620		
Performance	.733	.584	.690	.705	.800	.821	.790	.813		
Full Scale	.834	.684	.794	.776	.793	.804	.789	.765	.909	.952

next. The decision, of course, was necessarily made by the director of the project who accepts full responsibility for it. Various factors weighed on this decision.

1. Among the important ones, obviously, was the fact that any further attempts to approximate an ideal test would require that we petition the U. S. government for more time and money. To do this, it was felt that we would have had to offer strong evidence that material improvement could be achieved. While it seems that some further improvement could have been achieved, it appeared unlikely that Digit Symbol Substitution could have been. Illiterates just cannot do much with it. Also it does not appear that Picture Arrangement can be materially improved without substantially changing its character. This is also due to the illiteracy problem in the main. Neither do we see any way of improving Analogies (similarities). Hence it seemed that we were faced with justifying the use of the test as it is or recommending that we abandon the project.

2. Also considered was the fact that the Puerto Rican population probably is qualitatively different from that of the U. S. in that in Puerto Rico there is almost surely greater variance in the developed abilities of people. There are a substantial number of people in the Island who are as well educated and as capable as can be found in the States. In fact many of them were educated in the States. There are also a substantial number of people who have had little formal education and have had little opportunity to develop intellectually. Hence the problem of differentiation between people may well be more difficult. We require a test that is relatively overloaded with easy, primitive materials but which also carries enough difficult material to do a reasonable job in the higher levels as well. To try to do both with the same instrument threatened to require one which would have an undesirably long administration time, especially in view of our difficulty in identifying arbitrary starting points past item one of various subtests.

3. Since all aspects of the test function normally except for adequate discrimination between very low scores on certain subtests, the project director decided to compromise the difficulty problem somewhat and assemble the standardization form from the items used in the two major try-outs. It was clear that low scores could be reasonably well differentiated on any or all of the three total scores as, incidentally, could high scorers. Except for extremely low or extremely high scores, profiles will be as meaningful as they are in the States. In these cases substantial variations in profiles are rather unlikely to occur in any event. Of course, we also have the option of using the most advantageous age group for setting up the scaled scores. This

Table 16

FREQUENCY DISTRIBUTIONS OF THE VARIOUS SUBTESTS 124 SAMPLE

Score	Information	Analogies	Vocabulary	Digit Symbol	Picture Completion	Block Design	Picture Arrangement	Score	Arithmetic
100-103			0	0					
96-99			0	0					
92-95			0	0					
88-91			0	0					
84-87			2	0					
80-83			2	2					
76-79			6	2					
72-75			2	0					
68-71			2	2			0	17	2
64-67			5	3			0	16	1
60-63			6	3			1	15	2
56-59			7	4			1	14	5
52-55			5	9			7	13	8
48-51			7	6		0	9	12	8
44-47			9	6		3	5	11	10
40-43			9	13		3	11	10	15
36-39	1		9	9		6	14	9	9
32-35	4	0	13	8		9	9	8	14
28-31	12	1	13	9		13	5	7	21
24-27	15	8	11	7	10	12	12	6	15
20-23	16	14	4	10	19	18	10	5	8
16-19	23	19	5	9	39	13	8	4	3
12-15	32	27	2	6	23	19	10	3	2
8-11	17	24	3	5	20	6	12	2	0
4-7	4	25	2	3	11	16	8	1	0
0-3	0	6	1	8	2	6	4	0	1

seems to have some questionable features about it but if need be it can be done and that too will help to establish more meaningful profiles - at least for that age bracket. It almost surely would improve the apparent scalability of Digit Symbol considerably.

D. The Standardization Sample, N =1127, total population represented, statistical analyses and development of norms.

1. Means and Standard Deviations of raw scores by age groups.

The initial step was to compute means and standard deviations of the 11 subtests and the Verbal, Performance, and Full Scale Totals for the age groups presented in Table 17. The purpose of setting up this table was to get a closer look at age related trends in the scores than had been possible in the two smaller samples. Such trends might well determine what age band should be used as a reference group for setting up the conversion table for converting raw scores to scaled scores. For the U. S. WAIS the age band 20-34 was used because that was judged to be the period of maximum mental ability. It was the age span which, in general, produced the highest scores. In the U. S. case this was true of both the Verbal and Performance Totals and therefore of the Full Scale Total as well.

In the Puerto Rican population the Verbal and Performance Totals have rather different relationships with age. The Verbal Total means are rather flat for the most part. The 17 years old group has a relatively high mean, then the means appear to fall a little until age 30. The ages 30 to 49 tend to run somewhat higher than either the groups immediately younger or immediately older. The differences among the 13 groups, however, are not statistically significant (F=1.696; d.f.=12.1114). These data, then, do not provide any clear evidence of a relationship between age and Verbal Total. Because the over-all analysis of variance may have had poor power against it, it could still be that the two end groups are in fact different from the rest. Since a specific contrast involving this particular possibility was not pre-planned, it was not done. An a posteriori test of this kind would be inappropriate here.

In contrast to the Verbal Total, the Performance Total has a very strong relationship with age. It is highest for the 16 and 17 years old subjects and then slides downward rapidly to the lowest mean, that of the 60-64 age group. The mean of the latter is only about 44% as high as the mean of the former. This function is obviously highly reliable statistically.

It is also obvious that because of the Performance Total relationship the Full Scale Total also is related to age. Even so there is no clear relationship from about age 16 through age 44. This is due to a tendency for the higher verbal scores to occur in the age range of 30 to 49. Even though the apparent trend in the Verbal Total is not a reliable relationship, it still acts to negate, at least in part, the very clear Performance--age relationship.

2. Selection of the Reference Group.

From all the above it is clear that the reference group for scaling purposes could have been chosen from as wide a band of the subjects as ages 16 to 44. By the time one extends the reference group that far it would seem almost as reasonable to utilize the whole population as the basis for determining scaled scores. After consultation with the Psychological Corporation it was decided to use the same reference group they had -- ages 20 - 34. The argument here is that there is potential value in using the same reference group that has been used in the States and also in almost all other adaptations that have been made in Europe and elsewhere. To what extent such potential value will be realized is uncertain but at least such a choice in this case will do no clear violence to the scales to be developed for the Spanish WAIS and it will leave such cross cultural comparisons as possibilities.

The selection of the reference group, of course, might well have been decided on the basis of comparisons between the item statistics of various possible age bands. The only comparison of this kind that was actually made was between the 20 - 34 age band and the total sample. Item statistics for both will be presented in various of the sections that follow.

3. Uncontaminated item-subtest correlations for reference group and full sample.

The uncontaminated item-subtest correlations for the age 20-34 reference group and for the full sample are presented in Table 18. The correlations were all computed by correlating the item score with the subtest score minus the item score. These are, then, non-overlapping variables and hence contain no spurious correlation. The correlations are point-biserials which probably were more appropriate than biserials would have been since all of the subtest distributions were skewed to varying extents.

Inspection of the table shows that a large number of very high item-test correlations were obtained. In almost all cases in which a specific correlation is below .30 it is because the item was either very easy or very difficult. In either event the maximum possible correlation that can occur is severely limited. In the case of the Information and Arithmetic subtests, for instance, various of the

Table 17

PUERTO RICO - STANDARDIZATION SAMPLE N = 1127
MEANS AND STANDARD DEVIATIONS OF RAW SCORES

	Age	N	Inf.	Comp.	Arith.	Analog.	Dig. Sp.	Vocab.	Verbal Total
MEANS	15	17	11.94	8.06	7.76	7.47	8.00	21.94	64.88
	16	69	13.54	8.83	7.78	8.70	8.33	24.45	71.64
	17	66	14.67	10.70	7.64	9.62	8.44	29.53	80.59
	18-19	84	12.73	9.67	7.38	8.19	8.07	25.96	72.00
	20-24	152	13.26	10.29	7.49	9.02	7.80	28.38	76.24
	25-29	118	13.49	10.42	7.65	8.68	7.84	28.32	76.40
	30-34	106	13.81	11.62	8.08	9.48	7.94	30.08	81.01
	35-39	104	14.05	10.69	8.32	9.19	7.98	30.66	80.89
	40-44	127	14.93	11.29	8.61	9.86	7.91	32.50	85.10
	45-49	94	13.71	11.08	8.30	9.32	7.84	30.23	80.48
	50-54	62	13.06	10.82	8.00	8.98	7.79	28.94	77.59
	55-59	60	12.95	10.38	8.38	9.62	7.35	29.92	78.61
	60-64	68	11.66	9.35	7.41	7.72	6.75	25.15	68.04
	Possible Maximum Score		32	30	16	30	17	80	
STANDARD DEVIATIONS	15	17	5.06	3.28	2.58	4.30	1.72	12.12	25.75
	16	69	7.22	4.59	2.72	5.46	2.14	15.26	34.03
	17	66	6.14	4.38	2.63	5.29	1.71	13.39	29.30
	18-19	84	7.10	5.11	2.99	5.69	1.90	16.18	35.57
	20-24	152	6.76	5.21	3.16	5.74	2.16	15.73	35.33
	25-29	118	6.91	4.80	2.83	5.80	2.00	15.02	33.63
	30-34	106	7.04	5.32	3.05	5.85	2.30	15.97	35.90
	35-39	104	6.71	5.06	2.82	5.23	2.00	16.15	34.34
	40-44	127	7.84	5.59	3.48	6.25	2.36	18.44	40.56
	45-49	94	7.21	5.59	3.28	5.91	1.94	16.14	37.11
	50-54	62	6.07	5.24	2.80	5.56	1.68	13.75	31.54
	55-59	60	6.99	5.45	3.59	5.22	2.50	17.59	37.55
	60-64	68	6.01	4.57	3.05	4.55	2.20	15.46	32.21

Table 17
(Cont.)

	Age	Dig. Sp.	Pict. Comp.	Bl. Des.	Pict. Arr.	Obj. Assemb.	P	FS
MEANS	15	40.12	13.46	20.18	23.88	20.24	117.88	182.76
	16	42.14	14.61	22.93	24.48	19.77	123.91	195.55
	17	43.47	14.94	23.06	25.24	19.79	126.65	207.24
	18-19	39.33	13.86	21.08	22.61	18.39	115.27	187.40
	20-24	39.30	13.70	21.86	22.26	18.33	115.45	191.69
	25-29	36.56	13.29	19.97	20.79	16.93	107.54	183.94
	30-34	34.66	12.60	20.32	20.00	16.56	104.14	185.15
	35-39	31.25	12.04	19.81	17.77	16.11	96.98	177.87
	40-44	30.56	11.47	20.75	17.56	16.13	96.47	181.56
	45-49	22.90	9.82	18.78	15.12	13.86	80.48	160.98
	50-54	22.00	8.76	15.97	12.27	12.71	71.71	149.30
	55-59	19.63	9.47	16.62	12.70	13.10	71.52	150.13
	60-64	13.35	7.43	14.65	9.81	9.76	55.00	123.04
	Possible Maximum Score	110	25	46	48	44	273	478
STANDARD DEVIATIONS	15	7.95	3.90	8.91	10.35	7.28	28.49	48.89
	16	15.88	5.36	11.21	11.72	9.61	45.91	76.21
	17	14.36	5.52	10.37	10.71	8.92	40.05	64.72
	18-19	16.81	5.00	10.75	11.61	9.44	46.27	77.49
	20-24	17.54	5.07	11.46	11.66	9.76	48.15	80.11
	25-29	17.43	4.84	11.64	11.57	9.02	48.06	77.96
	30-34	18.57	6.19	10.97	12.19	9.79	50.38	83.30
	35-39	15.82	5.48	10.39	11.13	9.53	44.58	75.74
	40-44	20.47	6.39	11.60	11.94	10.59	55.26	92.67
	45-49	16.72	6.13	11.55	10.62	9.36	49.30	83.60
	50-54	16.28	5.45	10.13	9.91	8.34	43.35	72.13
	55-59	16.59	5.86	9.85	10.00	8.40	46.14	81.51
	60-64	11.89	5.51	10.10	8.52	6.73	37.69	94.08

Table 18

UNCONTAMINATED ITEM SUB-TEST CORRELATIONS FOR THE REFERENCE GROUP (N = 376) AND THE FULL STANDARDIZATION SAMPLE 1176

(Point Biserial Correlations)

	Information		Comprehension		Arithmetic		Analogies		Vocabulary	
ITEM	376	1127	376	1127	376	1127	376	1127	376	1127
1	0.086	0.088	.091	.080	.097	.137	.356	.346	.500	.530
2	0.151	0.145	.500	.480	.097	.111	.420	.371	.362	.393
3	0.000	0.073	.396	.413	.334	.295	.429	.436	.514	.508
4	0.202	0.164	.405	.390	.275	.253	.437	.417	.520	.524
5	0.229	0.211	.298	.342	.396	.379	.484	.449	.500	.532
6	0.428	0.434	.437	.419	.453	.439	.548	.508	.402	.431
7	0.459	0.486	.521	.518	.493	.527	.697	.650	.353	.394
8	0.339	0.356	.575	.551	.556	.570	.637	.611	.514	.553
9	0.409	0.407	.557	.605	.503	.534	.670	.650	.529	.506
10	0.425	0.432	.532	.530	.586	.606	.582	.549	.447	.484
11	0.571	0.580	.590	.586	.560	.581	.638	.659	.501	.507
12	0.601	0.603	.474	.455	.596	.582	.666	.679	.580	.549
13	0.585	0.566	.518	.491	.448	.498	.607	.598	.234	.274
14	0.564	0.575	.463	.489	.507	.501	.601	.583	.479	.526
15	0.682	0.682	.362	.375	.574	.559	.513	.492	.480	.543
16	0.703	0.685			.464	.451			.602	.585
17	0.647	0.664							.645	.634
18	0.646	0.639							.650	.646
19	0.557	0.591							.393	.448
20	0.574	0.621							.681	.656
21	0.700	0.680							.504	.543
22	0.534	0.545							.635	.579
23	0.731	0.745							.513	.481
24	0.599	0.580							.547	.560
25	0.624	0.593							.505	.477
26	0.512	0.524							.647	.651
27	0.649	0.620							.597	.610
28	0.577	0.549							.616	.640
29	0.632	0.600							.493	.527
30	0.618	0.586							.687	.673
31	0.517	0.532							.622	.637
32	0.461	0.548							.681	.686
33									.646	.646
34									.557	.584
35									.634	.638
36									.495	.536
37									.594	.650
38									.516	.525
39									.477	.586
40									.517	.598

Table 18
(Cont.)

	Picture Completion		Block Design		Picture Arrangement		Object Assembly	
ITEM	376	1127	376	1127	376	1127	376	1127
1	.347	.385	.365	.365	.396	.436	.570	.580
2	.253	.290	.379	.404	.487	.542	.580	.610
3	.308	.388	.542	.549	.597	.648	.539	.576
4	.391	.412	.660	.621	.609	.610	.492	.535
5	.424	.487	.615	.629	.622	.658		
6	.483	.490	.627	.653	.547	.527		
7	.478	.525	.654	.650	.601	.623		
8	.574	.606	.687	.658	.538	.566		
9	.503	.533	.454	.436	.645	.657		
10	.499	.601	.424	.401	.600	.599		
11	.444	.486						
12	.566	.636						
13	.622	.641						
14	.536	.622						
15	.487	.537						
16	.438	.511						
17	.418	.494						
18	.486	.531						
19	.392	.396						
20	.453	.574						
21	.499	.570						
22	.413	.469						
23	.470	.471						
24	.449	.441						
25	.285	.308						

low correlations which occurred on easy items were about the highest that could occur.

Item 13 of Vocabulary had a correlation with the sum of the other 39 items that was less than .30 but even this item is basically quite a good item. This was checked out by setting up a contingency table in which the subtest total scores were divided at the mean of the subtest. The 2, 1, and 0 scores for the subjects above the mean were 13, 111, and 49 and those below the mean were 3, 112, and 88 respectively. Clearly the item is contributing meaningfully to differentiation between upper and lower scoring persons. Contingency tables of this kind were set up for every item.

Finally, inspection of the table also shows a remarkable similarity between the correlations of the reference group and those of the full sample. Of course these correlations are bound to be somewhat similar due to overlapping of the sample. The closeness of the correlations, however, does give confidence that the age groups outside of the reference band of 20 to 34 are yielding substantially the same item-subtest correlations as the reference group.

4. Item difficulty and item order

The difficulty level of the individual items of all subtests except Object Assembly is defined as the percent of examinees who earned non-zero scores on the item. For such subtests as Vocabulary this meant that the number of examinees who earned scores of one was added to the number who earned scores of two and this sum was divided by the number in the sample. Argument might be made for using weighted sums for these subtests but this was not done. For the Object Assembly subtest the level of difficulty is defined as the percent of examinees who correctly assembled the entire object whether or not they earned time bonus.

The difficulty levels of each of the items are presented in Table 19. The difficulty levels are presented both for the reference group of 376 subjects who were in the 20 to 34 age band for the entire sample of 1127 persons.

It is clear that there are not many differences of much importance that are associated with the age span used for presenting these indices. Such differences as do occur are quite consistent with what one would expect from the fact that the verbal subtests show no clear time related trend and the performance subtests do show such trends.

It is also clear that the rank order of difficulty levels do not change much when one restricts the sample to the 20 to 34 age range. This is demonstrated with the correlations at the bottom of the table in which the rank of the difficulty level of the item is correlated with the order of the item in the standardization manual. If one correlated the two directly, these correlations would be still higher. The correlations chosen for inclusion here were considered essential to the decision which has to made as to whether or not the items should be reordered in setting up the final published manual. These correlations are high enough that we feel that the item should not be reordered. Since we do not have sufficient information concerning the possible effects of item sequence, the norms will be likely to be more exact if the item order remains as it was for the standardization testing.

Finally, and perhaps most important, it is clear from Table 19 that substantial spreads of item difficulty were achieved on all subtests. In some cases there were probably fewer very easy items than would be ideal. Vocabulary is a good example of this where 90 percent was the highest level reached. In other subtests, such as Block Design and Arithmetic, there are signficant gaps in the continuity of the difficulty levels. We cannot, therefore, hold that we have achieved as ideal a test content as we might have preferred. Doing so would have required one to two more years of efforts in the area of item development and testing.

5. Time bonus

Another characteristic of the test that had to be resolved concerned the assignment of time bonuses for quickness of responses to certain of the items of 4 of the subtests. In the English language WAIS time bonuses are awarded for some to all of the items on Arithmetic, Block Design, Picture Arrangement, and Object Assembly. A partial investigation of the effect of time bonus on item-test correlation and on reliabilities was carried out on all four of these tests using the 254 sample data and then replicated on the Arithmetic subtest in the Standardization sample.

It was clear from this analysis that Object Assembly was improved by incorporating time bonus. Item correlations with the subtest total minus the item were much higher -- the increase being from an average of about .20 to an average of over .50. For Picture Arrangement, awarding time bonus to about 5% of the responses had virtually no net effect on item versus subtest correlations. For Block Design awarding time bonus to about 5% of the responses did increase, on average, item-subtest correlations, although the increase only averaged about .04 or so.

The Arithmetic Subtest presented results that were somewhat different according to the age level of the subjects. When all age levels

Table 19

PERCENT OF SUBJECTS WHO PASSED INDIVIDUAL ITEMS OF EACH SUBTEST IN REFERENCE GROUP AND ENTIRE SAMPLE

Subtest Sample	Information		Comprehension		Arithmetic		Analogies		Vocabulary	
ITEM	376	1127	376	1127	376	1127	376	1127	376	1127
1	100	100	94	97	100	100	87	86	82	79
2	99	99	83	83	100	100	91	92	90	88
3	100	100	93	91	94	94	77	73	85	84
4	96	97	71	72	93	94	86	84	90	89
5	94	95	48	50	81	82	80	80	79	80
6	80	77	71	69	76	80	50	51	82	83
7	77	74	40	40	44	46	53	55	59	62
8	49	52	48	46	49	52	48	51	78	78
9	60	63	37	39	26	28	53	50	85	83
10	57	55	55	49	26	31	32	31	70	74
11	41	47	48	44	23	26	23	22	80	81
12	48	46	27	29	22	21	22	24	67	69
13	37	37	21	19	7	10	16	19	64	68
14	54	52	15	16	11	11	14	12	72	74
15	45	42	6	8	12	13	8	9	69	68
16	34	32			6	6			62	60
17	34	35							59	58
18	22	22							37	36
19	18	21							63	61
20	19	26							51	48
21	23	21							44	44
22	21	23							62	58
23	24	25							51	49
24	29	25							41	46
25	15	14							32	39
26	9	11							33	33
27	14	14							32	31
28	14	12							30	30
29	12	11							23	24
30	11	10							30	25
31	7	8							28	27
32	6	9							23	25
33									30	25
34									23	21
35									22	21
36									15	17
37									15	15
38									11	10
39									9	13
40									10	12
Correlation of difficulty order with standardization order	.970	.979	.931	.946	.984	.977	.968	.971	.963	.975

6

Table 19
(Cont.)

Subtest Sample	Picture Completion		Block Design		Picture Arrangement		Object Assembly	
ITEM	376	1127	376	1127	376	1127	376	1127
1	85	80	95	94	97	95	63	60
2	97	96	92	91	95	93	39	33
3	92	87	71	70	61	55	34	29
4	81	80	61	62	68	61	22	21
5	88	81	54	50	53	48		
6	84	74	53	49	36	30		
7	69	62	39	36	36	29		
8	76	71	33	29	27	23		
9	65	62	11	9	37	31		
10	71	61	8	6	30	26		
11	56	54						
12	61	51						
13	51	40						
14	63	54						
15	52	43						
16	36	33						
17	40	34						
18	29	27						
19	19	17						
20	30	26						
21	26	23						
22	18	17						
23	17	14						
24	14	16						
25	7	6						
Correlation of difficulty order with standardization order	.999	.999	1.00	1.00	.89	.89	1.0	1.0

were included the item-subtest correlations tended to decrease among those items not receiving time bonus and to increase among those items receiving time bonus. The net effect here was more or less balanced. Reliabilities were also essentially uneffected in the total sample. However, if one subdivides the sample into younger and older groups, then he finds that the reliability of the younger groups drop slightly when time bonus is involved and remain stable or rise slightly for the older age span.

A special comparison was set up in the Standardization sample in which the Arithmetic scores of the age 20-24 age group was compared with those of the 55-64 age group, the comparison being made with and without time bonuses being involved. Amazingly, the old group earned somewhat more time bonus (although this may not have been reliably more in the statistical sense). This means that awarding time bonus on Arithmetic would have tended to increase slightly the rising trend of Arithmetic Scores as age increases. The older group also had slightly better reliability of performance than the younger when time bonus was involved. The reliability of the older group remained essentially constant with time bonus (.901 with and .908 without) whereas the reliability of the younger group dropped slightly (from .883 without time bonus to .868 with it).

Since the test is almost surely going to be used much more extensively among younger people and since no clear advantage and possibly some disadvantage in using time bonuses accrue to the young group, it was decided not to include such bonuses in the Arithmetic Subtest.

Since there is clear advantage in using time bonuses for Object Assembly, probably a little advantage in using them for Block Design, and probably no disadvantage in using them for Picture Arrangement, such bonuses were left in the scoring keys for these subtests. It is expected that this problem of the role of time bonuses in test performances will eventually be studied intensively and will form the basis for a journal publication.

It is debatable as to how much of the total variance of a test of intelligence should be accounted for by a time bonus factor. It is estimated that the contribution of the time bonuses to the variance of the full scale is only about 4%. There is none in the verbal total.

6. Derivation of the Scaled Scores

Having resolved the questions of item content, time bonuses, and of the reference group on which scaling was to be carried out, the next step was to derive Scales Scores. These Scaled Scores were to be such that they would form normal distributions with means of 10 and standard deviations of 3.

The procedure for scaling is described in part in Chapter 2 of the manual which appears as Appendix I to this report. It consists of setting up cumulative frequency distributions of the raw scores, setting each cumulative percentage at its corresponding standard score value on the theoretical unit normal curve, and then multiplying the resulting standard score by three and adding ten. The end points are spread slightly by using the ratio of the y-ordinate to the area under the curve from the end proportion out toward the tail instead of standard scores for these two points. All points in between are stabilized to some degree by using a smoothing procedure which involves adding succeeding frequencies together and then setting up the cumulative distribution with these combined frequencies.

The actual process of scaling was accomplished without difficulties of note. One element in it should be mentioned. In order to maximize the differentiating properties of the low scores by making full use of all possible scaled scores values, it was necessary to leave a slight degree of negative skewness in some 5 or 6 of the subtest distributions. As will be seen later, this had very little effect on the full scale distributions. Tests of significance of deviation from the theoretical normal curve either for the subtests or for the total scores would fail to detect this unless sample sizes were very large -- probably as much as 500 or more would be required. This statement, of course, holds only for the reference group. Some skewness would have occurred in other age bands even without this having been deliberately left as a slight effect in the reference group. The skewness in other age bands may have been slightly enhanced by it.

The scale score conversions are presented in Table 15 of the manual which is included as Appendix I of this report. These conversions were actually carried out on all protocols in the standardization sample. Having them one can immediately answer the question as to how much difference there would be between the means and standard deviations of the 20-34 age group and the means and standard deviations of the full sample. The means and standard deviations of both are presented in Table 20.

The difference that occur between the means and standard deviations of the 6 verbal subtests are trivial. In-so-far-as these 6 subtests are concerned either the age 20-34 group or the full sample could have been used as the reference group. The means of the 5 performance subtests are all lower in the full sample than they are in the 20-34 age groups while the standard deviations are

Table 20

MEANS AND STANDARD DEVIATIONS OF THE 14 TEST SCORE VARIABLES FOR THE 20-34 REFERENCE GROUP (N = 376) AND THE FULL SAMPLE (N = 1127)

	20-34 Reference Group		Full Sample	
Variable	Mean	S.D.	Mean	S.D.
1	10.03	3.05	10.05	3.06
2	9.93	2.93	9.78	2.98
3	10.08	3.07	10.29	3.02
4	9.97	3.22	10.00	3.11
5	10.03	2.96	10.01	2.92
6	9.95	3.13	9.95	3.23
7	9.93	3.16	9.09	3.41
8	9.93	3.02	9.28	3.36
9	9.88	3.14	9.63	3.10
10	9.87	3.10	9.28	3.28
11	9.96	3.00	9.58	3.12
12	59.99	15.79	60.08	15.58
13	49.58	13.35	46.48	14.25
14	109.57	27.90	106.94	28.09

generally quite similar. If for the sake of statistical argument one considers the 20-34 age group to be a random sample of 376 from a population of 1127 (the full sample) and then calculates t-ratios adjusted for proportion of overlap he finds that Digit Span (t = 5.37), Picture Completion (t = 4.22), Picture Arrangement (t = 3.24) and Object Assembly (t = 2.66) all have reference means that are reliably higher than the full sample means. Only Block Design (t = 1.76) was not reliably higher. These differences are due in part, but only in part, to the fact that the younger groups earned more time bonus than the older.

It is clear that using the entire sample as the reference group for deriving Scaled Scores would have produced a different outcome. The differences would have been observed primarily in the Performance Subtests, the Performance Total, and the Full Scale Score. In the former cases the mean of the younger groups would have tended to be higher than 10.0 and the mean of the older groups below 10.0. For the Performance Total the mean of the younger group would have been somewhat above 50 and the means of the older would have been somewhat below 50. In all cases, considered collectively, there might have been a little better balance in that deviations in both directions might make for the least net deviations of age group means from the reference mean.

7. IQ equivalents of sums of Scaled Scores.

The transformation of the sums of the Verbal, Performance, and Full Scale Totals of the Scaled Scores was carried out with a simple linear transformation of the form $Y = aX + b$. This, of course, results in the form of the distribution of the IQ equivalents for any given age band being exactly the same as that for the scaled score distribution. That is to say, even for the 20-34 age band there will be a slight skewing. The actual amount can be illustrated by the fact that the means and medians of the three scale score sums are for the most part within one or two points of each other.

The IQ equivalents of the three sums of Scaled Scores for each of 6 age bands are presented in Table 16 of the manual which is Appendix I of this report (see pages 104 to 115 of Appendix I).

Perhaps it should be pointed out that ages 16 to 19 were combined into one age band for constructing the tables of norms presented in Table 16. This was because we did not obtain any clear evidence that the 16 and 17 year olds were different in performance from the 18 and 19 years old. There was some suggestion that the younger half of this group might actually have slightly higher mean Scale Scores than the 18-19 year olds. Their full scale, scale score mean was 4.72 points higher but the t-ratio for this difference was only 1.32. Besides that fact is the fact that the 18-19 year olds group reported an average of nearly 1/2 year more of education. With that much more education one would expect the older group to have the higher mean test performance. The younger examinees had a higher scale score verbal total by 2.16 points and a higher scale score performance total by 2.55 points. Neither of these differences were significant either, the t-ratios being 1.03 and 1.53 respectively. Since the evidence was at best inconsistent it was decided to simply pool all examinees in this entire age span. It is clear that very little error in estimating an IQ will occur because of a person's age within this range.

8. Intercorrelations among the 11 subtests and of the 11 subtests with the Verbal, Performance, and Full Scale totals.

These correlations are presented in Tables 7, 8, and 9 of Chapter II of the manual which is included as Appendix I of this report. They are presented for the age standardization groups 16-19, 25-34, and 45-54. All subtests are substantially intercorrelated. The 6 Verbal Subtests are intercorrelated among themselves and with the Verbal Total to a greater extent than they are with the 5 Performance Subtests and the Performance Total. The 5 Performance Subtests in turn are more highly intercorrelated among themselves and the Performance Total than they are with the Verbal Subtests and the Verbal Total. All 11 subtests are correlated quite highly with the Full Scale Score.

A comparison of all these correlations with the corresponding correlations obtained in the U. S. standardization shows that there may be some tendency for them to be slightly higher, on average, than the U. S. correlations. To the extent that this is true it seems to be more true of the Performance Subtests than of the Verbal. There also seems to be something of an age related trend among the Puerto Rican correlations that is not apparent in the U. S. data, and, again, this appears to be stronger in the Performance Subtests than in the Verbal, Over-all it is clear that the Puerto Rican results are no less favorable than the U. S. results.

9. Reliabilities of the 11 subtests and of the Verbal, Performance, and Full Scale Totals.

The reliabilities of the 14 scores derived from the spanish WAIS, along with their standard errors of measurement, are presented for the age groups 16-19, 25-34, and 45-54 in Table 6 of the manual presented in Appendix I. The reliabilities, except for Digit Symbol, are all internal consistency estimates. Digit Symbol is a test-retest reliability on the same form. It is our opinion that test-retest reliabilities are on various counts more realistic and for this reason we had hoped to have the time and resources to select a subset of the standardization sample and retest them. Time and resources had to be diverted to the unexpected item development phase which resulted in the 124 sample testing and item analyses. Hence the reliabilities had to be estimated through split-half (i. e., internal consistency) methods. The same methods used by the Psychological Corporation for the U. S. analysis were employed for this analysis. This does have the advantage of allowing direct comparison of the level of reliability achieved in the Puerto Rican sample with the level of reliability achieved in the United States.

Comparison of Table 6 in the appendix of this report with Table 6 in the U. S. WAIS manual will show that for all three age levels the reliabilities of the Puerto Rican data are a little higher for almost every subtest and for the Verbal, Performance and Full Scale Scores. It should be added that the reliabilities estimated for the Performance and Full Scale Scores were deliberately made conservative by entering estimates for Digit Span into the formula which were either age regressed (for the oldest group) or treated as if they had been only one half as long (for the two younger groups). It should also be pointed out that the higher reliabilities were obtained in the Puerto Rican data even though two (the 25-34 and 45-54) age groups had smaller sample sizes. It is abundantly clear that on the dimension of reliability and therefore of Standard Errors of Measurement the Puerto Rican WAIS compares quite favorably with the U. S. WAIS.

10. Concurrent Validity.

The only form in which the potential validity of the Puerto Rican WAIS was,investigated directly was in effect, to develop a second Intelligence scale. This was accomplished by utilizing a Spanish Language form of the Otis Higher Examination, Form A which had been prepared by Lcdo. Robert B. MacVean, Director, The American School, Guatemala City, Guatemala, and his research associates.

We had every person who had had 7 years or more of education and who was in the second half of the standardization sample take the Otis as the 12th subtest in the examination procedure. The

Otis is a timed test so this added 30 minutes to the testing sessions of these examinees. A total of 205 such persons took the Otis. This was an island wide sample, representative of the total population between 16 and 64 who had completed 7 or more years of education. There may have been a small bias in the Urban-Rural residence ratio in this sample but this possibility was not specifically traced out.

An item analysis of the Otis data was carried out. Since the Otis form was translated in Guatemala, and since this was done without including an overt attempt to be sure that the form would be as widely applicable as could be, it was to be expected that some items might not be satisfactory. This result did, in fact, happen. Over-all, however, it is clear that the Otis functioned quite well in Puerto Rico. Of the 75 items on the test 57 had uncontaminated point-biserial correlations with the total score ranging from .300 to .605. Among the remaining items there were 5 in the interval .250 to .299, eight in the interval .200 to .249, three in the interval .138 to .199 and finally only two in the interval .00 to .137. This last interval represents those items which failed to show a statistically reliable positive item-test correlation (where alpha is considered to be .025).

The Otis Score correlated .856 with the WAIS Verbal Score, .725 with the WAIS Performance Score, and .853 with the WAIS Full Scale Score. Since its content is clearly more similar to the content of the Verbal Subtests of the WAIS than to the Performance Subtests, this outcome represents a clear demonstration of concurrent validity for the Spanish WAIS.

There was one outcome in the data from these 205 subjects that is not directly related to the Otis outcome that should be mentioned. Among these subjects who have had 7 years or more of education, the correlation between the Verbal Total and the Performance Total is lower than it is when the full range of education is included (.694 versus .810) and somewhat more surprisingly the Performance Total correlates a little more highly with the Full Scale Score than the Verbal does (.907 versus .933). This last difference is in the opposite direction from what is almost always observed in any sample. It raises the possibility that as education increases, there is a tendency for it to have relatively greater effects on functions involved in the Performance Scales than it has on the Verbal Scales.

Since the Otis was speeded and since no retest data was available, no estimate of the reliability of the Otis was calculated. It is clear that the lower bound of its reliability would be the .856 correlation between it and the Verbal Score of the

WAIS. Since the group on which this correlation occurred was restricted on the education variable, this lower bound for reliability indicates that the test can be used at least as a preliminary evaluation of mental ability with considerable confidence. The upper limit of its standard error of measurement is 4.65 which compares favorably with this statistics when obtained from short forms such as the Otis in the United States.

11. An Optimum short form of the Spanish WAIS.

If for some reason one wishes to use a subset of the WAIS subtests rather than the Full Scale to estimate an IQ then he has the problem as to which subset would be in some sense optimum. For the guidance of such persons a stepwise multiple linear regression analysis, developed by M. A. Efroymsen, was carried out on the 376 reference group cases. This method consists of calculating an initial correlation matrix of the 11 subtests and the Full Scale Score as a 12th variable. The subtest which correlates highest with the Full Scale Score is partialled out of the matrix and then the subtest which correlates highest in this reduced matrix is partialled out and so on until as much of the variance has been accounted for as one wishes. It might still be that some other set of 4, or whatever number of variables one might wish to use, will correlate higher than the set to be presented here. It is probable, however, that such a set would do so because of capitalizing on chance elements in the correlations rather than this representing an optimum accounting for the true variance of the Full Scale Scores.

Vocabulary was the best single subtest predictor of the Full Scale Scores (R = .886). The best pair of subtests proved to be Vocabulary and Picture Completion (R = .94). The best triad was Vocabulary, Picture Completion, and Arithmetic (R = .964). The best four subtests were Vocabulary, Picture Completion, Arithmetic, and Picture Arrangement (R = .976). Longer sets would add, in order, Block Design (R = .983), Analogies (R = .988), Digit Symbol (R = .992), and Comprehension (R = .994). The "best" four subtests account for 95.2% of the variance of the Full Scale Scores and so would do a pretty good job of estimating it. If all one wanted was a score which would be a reasonable estimate of mental ability, then a shortened version of the test could reasonably be used. We would like to add that we agree with Wechsler's comment in this regard that one gives individual tests in part because it affords a special kind of situation in which much more than a score can be obtained. It provides an excellent opportunity to observe and gain impressions of many other characteristics of an examinee as well. For the sake of obtaining these observations as reliably as possible one should use the whole test unless the situation compells settling for less.

E. The problem of predicting the rate of change of IQ and the need for automatic adjustment of norms or restandardization.

The contract for this project called for making an effort to predict the rate of change of IQ scores in the population. This was deemed necessary because the number of years being spent in school and the quality of the schools are increasing rapidly in Puerto Rico. Since scaled scores and therefore IQ scores are substantially correlated with number of years in school, increases in the amount and quality of education are bound to cause the norms developed and presented with this report to become inaccurate in a relatively short time.

The problem was approached by obtaining extra cases for carrying out two additional kinds of analyses that could not be carried out satisfactorily on the standardization data itself. The first of these was a set of randomly selected cases to augment the number of persons with little education in the 25 to 29 age group and the number of persons with higher educational background among examinees in the 55 to 64 age span.

1. The analysis of change in scores with age when education is balanced.

Four age groups (25-29, 35-39, 45-49, and 55-64) were formed by combining the special cases tested with the subjects in the appropriate standardization age groups in such a manner as to provide almost exactly the same total number of years of education in each group. This required that some people with 12th grade and higher education be randomly dropped from the younger groups and that some people with five years and less be randomly dropped from the older groups. The initial composition of the total groups, the numbers dropped from each by random methods, and the final composition of the groups are all presented in Table 21. In reviewing Table 21 the reader should keep in mind the fact that the two year intervals representing the numbers of years of education completed go through the whole year and not to the half year. In other words the 0-1 interval means 0-1.999 and therefore has an average of 1.

From Table 21 it is clear that the educational history of the groups has been balanced about as well as is reasonable to hope for. Of course, the means of the groups vary slightly more than is indicated here because they were computed from the grouped frequencies presented. The exact means are very closely represented by these numbers. It is important to state this because one of the most surprising and important results of this whole research project depends on its validity.

Table 21

THE ORIGINAL AUGMENTED AGE GROUPS, THE NUMBERS DROPPED FROM EACH BY RANDOM METHODS AND THE COMPOSITION OF THE EDUCATIONALLY BALANCED GROUPS

Years of Education Completed	Augmented Age Groups				Numbers Dropped				Numbers in Education balanced groups			
	25-29	35-39	45-49	55-64	1	2	3	4	1	2	3	4
18+	1	1	1	1	0	0	0	0	1	1	1	1
16-17	9	7	7	7	2	0	0	0	7	7	7	7
14-15	8	8	6	4	2	2	0	0	6	6	6	4
12-13	34	28	23	10	21	14	10	0	13	14	13	10
10-11	10	9	10	13	0	0	0	0	10	9	10	13
8-9	25	25	25	25	0	0	0	0	25	25	25	25
6-7	18	15	16	16	2	0	0	0	16	15	16	16
4-5	31	33	31	40	0	0	0	8	31	33	31	32
2-3	11	23	21	22	0	12	10	11	11	11	11	11
0-1	15	17	15	18	0	2	0	3	15	15	15	15
	162	166	155	156	27	30	20	122	135	136	135	134
Mean									7.76	7.74	7.76	7.58

One other factor in the balancing of the groups should be mentioned before looking at the mean performances of each. In order to keep the groups large and to achieve balancing that would leave almost no variability between the mean number of years of school, it was necessary to survey the adjoining standardization age groups for a few additional cases. That is to say, the 25-29 group has 15 subjects in it who were 24 or 30. The 35-39 age group has 19 who were 34 or 40. The 45-49 age group has 9 who were 44 or 50 and the 55-64 age group has 8 who were 53 or 54. At the younger levels this meant adding primarily low education people and at the older levels this meant adding primarily high education people. The mean ages of the groups were essentially unchanged by this augmenting procedure (except for the oldest group which on average was made slightly younger).

With the above background in mind let us look at the means and standard deviations of the resulting education balanced age groups. These results are presented in Table 22.

The Verbal Total increases significantly with age (F = 10.42, d.f. = 3.536)! This is contrary to the expectation of most investigators who have written about and studied age related trends.

The Performance total does not show a statistically reliable time related trend. All our expectations are in the direction of a decline in the Performance total with age.

The Full Scale Means of the Scaled Scores do not show a statistically reliable time related trend! The numerical values of the Performance Total do run slightly counter to the significant trend among the Verbal Scores in that they show slight, even though non-significant, decline in the two oldest groups.

These results are especially surprising in that the average education of each group averaged only about 7 and three fourth years. Possible biases in the form of the selection processes that might have operated in determining which of the people would get to go to school in past decades come to mind immediately. Against such biases are the facts that those who did go went to poorer schools than are now available, under more poorly trained teachers, and were mostly forced to use a "foreign" language while in school -- i. e., English. In any event an average of seven and three quarters years hardly constitutes high level education.

Neither can one appeal to the proposition that including the whole education spectrum in the groups accounts for the results. If one subdivides the subjects and obtains means on those with 6

Table 22

SCALED SCORE MEANS AND STANDARD DEVIATIONS OF 4 AGE GROUPS BALANCED IN AMOUNT OF EDUCATION

	Age Group							
	25-29		35-39		45-49		55-64	
Test	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Information	9.38	3.04	10.35	2.80	10.90	2.97	11.07	3.15
Comprehension	9.27	2.80	10.18	2.68	10.62	3.22	10.97	3.08
Arithmetic	9.49	2.97	10.57	2.67	11.21	2.79	11.17	3.26
Analogies	9.34	3.23	10.32	2.72	10.64	2.97	10.78	3.01
Digit Span	9.35	2.88	10.15	2.82	10.54	2.91	10.05	2.84
Vocabulary	9.27	3.28	10.45	2.90	10.87	3.32	11.38	3.26
Digit Symbol	9.10	3.08	9.10	2.97	8.34	2.85	7.84	3.08
Picture Completion	9.24	2.76	9.34	3.04	8.73	3.31	8.65	3.38
Block Design	9.01	3.24	9.68	2.90	9.76	3.02	9.55	3.00
Picture Arrangement	9.12	2.86	9.32	3.03	8.78	2.61	8.44	3.06
Object Assembly	9.45	2.94	9.67	2.84	9.31	3.09	8.84	2.96
Verbal Total	56.10	15.46	62.02	13.83	64.79	15.61	65.43	16.03
Performance Total	45.92	13.02	47.11	12.57	44.91	12.98	43.32	13.70
Full Scale Total	102.01	27.35	109.13	24.97	109.70	27.39	108.75	28.56

years of education or more separately from those with 5 years or less the same results are obtained although there is more of a suggestion of time related trends in the Performance Totals of the low education groups and less so in the higher education groups.

Inspection of the individual subtest results shown in Table 22 reveals other surprising results as well. For instance, Arithmetic and Digit Span are usually expected to decline because of a supposed connection with a immediate memory ability that is supposed to decline from about age 30 on. They do not decline in these education balanced groups.

Not all of the results run counter to expectation. Digit-Symbol Substitution, for instance, is supposed to decline with age because of a general slowing of physical responses. Digit-Symbol does in fact show a significant decline with age, primarily in the two older groups (F = 5.70; d.f. = 3.536). Other Performance subtest do no show clear time related trends.

It would appear that by far the dominant influence in age related trends is that Education Variable, at least for the ages from 25 to 64.[1] So far as the analyses of this phase of the data have gone, no sign of any other conclusion has appeared. This being the case the next step can reasonably be to attempt to estimate to what extent scores vary with education.

2. The estimation of changes in scaled scores as a function of education.

This problem was approached three ways. The first was to obtain the exact mean education of each age group in the standardization sample. These means and the Verbal, Performance, and Full Scale Means and Standard Deviations are presented in Table 23.

Across the ages 16-30 there is neither an apparent trend in the number of years of education nor an apparent trend in the Full Scale Scores of the groups. Actually the numerically, but not reliably, highest Full Scale Mean was obtained by the 16-17 age group (which includes 17 examinees who were 15 years of age). From age 30 on through age 64 there is no time or education trend related to the Verbal Total. (Recall, however, that for education balanced groups, the Verbal Total rises at least to age 40 and perhaps even beyond). Time related trends do occur in the Performance Total and because of this in the Full Scale Scores. It would appear that at least from about age 30 on, the declines in scores may well be completely accounted for by the decline in the education variable. This is supported by the fact that the 35-39 and 40-44 age groups have almost identical educational backgrounds and also have almost identicle mean scores on all three variables. The phenomenon occurs again that the 50-54 age group has almost identically the same educational level as the 55-59 age group and both of these also have almost identicle mean scores on the three variables.

---

[1] This finding is being subjected to a much more intensive investigation the results, of which, will be published in a Psychological Journal as soon as possible.

If one agrees that the age span from 16 to 30 has no meaningful trend and looks at the remainder of the age groups, it appears that the Full Scale Total is rising at a rate of about 5 score units per year of education. It was decided to look at this possibility in another way. The subjects in the age range 55 - 64 were sorted out and the remaining 997 subjects were sorted according to the number of years of education they had had. Ages were allowed to vary in the groups. The Means and Standard Deviations of the examinees appearing in each successive 2 years of education were then computed and are presented in Table 24.

It is obvious that all three scores rise steadily as education rises. In fact one can fit a linear equation to all three relationships quite satisfactorily. The Verbal Total rises at a rate of almost exactly three points per year of education. The Performance Total rises at a rate of almost exactly two points per year, and therefore the Full Scale Score rises at a rate of almost exactly 5 points per year of education. This result compares very closely with the one above in which the rate of change was also estimated at about 5 points per year of education. This result also supports the general conclusion that most if not all of any age related variance that occurs in the standardization sample can be accounted for primarily in terms of the educational background of the examinees. One might have thought that since younger examinees tend to be found in somewhat higher proportions in the high educational groups and the older examinees in the lower educational groups, the number of points estimated to be associated with a year of education might have been a little less in this case than it was before. In this case, however, the examinees in the two oldest 5 year periods were omitted and the younger people outnumbered the older by a ratio of about 2 to 1. These two factors may have been sufficient to prevent any change of this kind. Another factor involved might be the improving quality of the schools that is taking place.

Finally, the above conclusions could also be checked by computing the means and standard deviations of the three total scores of

Table 23

MEANS AND STANDARD DEVIATIONS OF THE 3 TOTAL SCORES OF ELEVEN AGE SUBDIVISIONS AND MEAN EDUCATION OF EACH GROUP

Age	16-17		18-19		20-24		25-29	
N	152		84		152		118	
Variables	M	S.D.	M	S.D.	M	S.D.	M	S.D.
1. Verbal Total	59.65	14.08	57.49	16.01	58.89	16.32	59.72	14.90
2. Performance	53.66	11.20	51.11	12.77	50.84	13.17	49.23	12.62
3. Full Scale	113.32	23.59	108.60	25.58	109.72	28.40	108.45	26.34
4. Education	8.49		8.46		8.43		8.36	

	30-34		35-39		40-44		45-49		50-54	
	106		104		126		94		62	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
1.	61.88	15.80	61.78	13.97	62.71	17.55	61.14	15.45	60.53	13.18
2.	48.16	14.20	46.35	12.28	45.89	15.76	41.53	13.86	39.50	12.30
3.	110.04	28.85	108.13	24.90	108.60	32.25	102.67	28.12	100.03	24.24
4.	7.54		6.82		6.81		5.64		4.60	

	55-59		60-64	
	60		69	
	M	S.D.	M	S.D.
1.	60.05	16.31	55.54	15.59
2.	39.68	13.74	34.57	13.30
3.	99.73	29.28	90.10	27.81
4.	4.50		3.48	

Table 24

MEANS AND STANDARD DEVIATIONS OF THE THREE TOTAL SCORES WHEN THE STANDARDIZATION SAMPLE IS GROUPED ACCORDING TO EDUCATION

N	Years of Education	Exact Average of Education of Group	Verbal M	Verbal S.D.	Performance M	Performance S.D.	Full Scale M	Full Scale S.D.
39	16+	16.87	86.90	11.73	65.97	10.40	152.87	20.50
36	14-15	14.33	83.14	7.81	64.19	7.81	147.33	13.34
171	12-13	12.13	72.22	10.98	58.68	8.88	130.90	17.93
120	10-11	10.48	66.40	9.55	56.10	9.30	122.50	16.57
144	8-9	8.45	60.75	10.07	50.70	9.36	111.45	16.53
117	6-7	6.35	55.17	10.51	45.59	8.92	100.76	16.46
157	4-5	4.51	52.45	11.37	41.68	8.59	94.13	17.02
128	2-3	2.63	50.69	9.26	37.06	9.95	87.75	16.90
85	0-1	.35	41.76	11.24	28.11	8.33	69.87	17.47

the parent-child data which was collected. Another set of supplementary cases was obtained by testing a parent of a random set of the children who were in the standardization sample and between the ages of 16 and 25 on the one hand and also a random set of children of parents who were in the standardization sample and who had children still living at home who were between the ages of 16 and 25. The means and standard deviations of these parent-child pairings, divided into three age groups according to the ages of the parents, are presented in Table 25.

The overriding influence of education on scoring is again illustrated in Table 25. Just about half of the parent in each group were also in the standardization group. This means that for the youngest parent group 17.8% of them were in the standardization group, for the middle group the overlap was 28.8%, and for the oldest group the overlap was 24.8%. There was, then, a considerable degree of independence between the parent-child data and the standardization sample data. This is important because of the pattern of outcome to be described. The youngest set of parents had an average of 1.1 years less education than the comparable age group in the standardization sample (5.71 years versus 6.82). The Full Scale mean of this group was 7.68 points lower than that of the standardization group (100.60 versus 108.38). This difference in scoring is significant in the statistical sense (t-2.1). The mean education of the other two parent groups almost exactly equalled the average education of the two comparable standardization groups and all three of their total scores were also almost identical to the means of the standardization groups.

As is noted in Table 25 the average rate of change of Full Scale Scores of children as compared to their parents was 3.56 points per years of education. This result was clearly independent of the age differences between the parents and their children. In fact the largest, although probably unreliably so, rate of change within one group occurred in the parent-child pair having the smallest age difference. It seems likely that this estimate of the rate of change associated with education is less than those mentioned earlier because of the greater communality of environment shared by the parent-child groups as compared with the independent groups which produced the estimate of 5 points.

What conclusion concerning adjutsments in norms can be justified? It is clear that for people who were 25 years of age or more in 1965 the age used for entering the IQ conversion tables should be fixed at the age they were in 1965. In other words the examiner should determine the birthday of the examinee, use this to determine how old he was in 1965 and then, if he was 25 or over at that time, enter the norm table which includes that particular age.

Table 25

MEANS AND STANDARD DEVIATIONS OF THE THREE TOTAL SCORES AND MEANS OF AGES AND YEARS OF EDUCATION OF PARENT CHILD PAIRS

	Age Band of Parents					
Parents	35-44		45-54		55-64	
N	73		97		60	
	M	S.D.	M	S.D.	M	S.D.
Verbal	57.80	14.29	61.04	14.17	58.93	13.23
Performance	42.81	12.18	40.79	12.46	36.65	12.56
Full Scale	100.60	25.27	101.83	25.46	95.58	24.70
Age	40.03		49.24		59.00	
Education	5.71		5.39		4.05	
Children						
Verbal	60.68	15.52	60.16	16.15	61.45	16.94
Performance	55.23	12.56	53.43	12.32	52.85	13.65
Full Scale	115.92	26.69	113.59	27.46	114.30	29.63
Age	18.03		19.63		21.43	
Education	9.29		9.26		9.33	
Diff. in Education	3.58		3.87		5.33	
Diff. in Full Scale Scores	15.32		11.76		18.78	
Rate of Change	4.28		3.04		3.52	

Average rate of change: 3.56 Full Scale Points per 1 year of difference.

For people who are under twenty-five at the time of testing or who were under twenty-five as of 1965 the answer is more difficult. For these age spans adjustments should be made according to the average number of years of education of the age group. This can be estimated from what is known about the rate of change of average education in the population. Some statistics relevant to this question are presented in Table 26. As can be seen in Table 26, the average education achieved by each succeeding age bracket has been increasing by about .75 years for each 5 year period for the past 30 years or more. That trend may actually have accelerated somewhat during the past 5 years. A higher rate of gain, however, probably cannot be maintained too far into the future. The youngest age group can be expected to reach a number of years saturation point within the next 5 to 10 years. This means that within about ten years the Full Scale norms of the 16 to 19 age group probably should be adjusted such that a scaled score of some 12 points less than now should be assigned a mean IQ of 100. The 20 to 24 and 25 to 29 age groups probably will need adjustments at about the same time. For the 20 to 24 group the adjustment probably will be such that a Scaled Score of some 5 points less will be assigned a mean IQ of 100. The 25 to 29 adjustment will probably be about 10 points. Just when these adjustments are made and how large they are should be determined from the educational statistics on mean number of years in school for each age kept by the Department of Education of the Island. A review of tables 22-25 indicates that no adjustment in the Standard Deviations is likely to be needed. It should be feasible to simply shift the whole distribution for each age band by an appropriate constant which will be the difference between the average number of years of education of the age 1965 standardization and the average for that age at the time the adjustment is made multiplied by five.

After about 15 years a random sample of about 250 persons should be drawn from ages 16-45 to make a more exact determination of the norms that should be in use.

F. The prediction of norms from a relatively small sample.

Since it is hoped that the test developed in Puerto Rico will be adopted for use in other Spanish speaking areas, and since it is to be expected that local differences in language usage educational level, and customs would require some adjustment in norms and scoring keys, ways of facilitating its economic adaptation to other areas are needed. To the end of showing that these needs can be met, it had been intended that the data from the 254 item analysis sample would be used to set up predicted norms to compare with the eventual

Table 26

RATE OF CHANGE OF EDUCATION IN THE POPULATION

Age at time of 1960 Census	*Median* Education of Age Group in 1960	Age at Time of Testing in 1965	*Mean* Education Reported in 1965	Maximum Mean which might be achieved by the age group at some future date
		16-19	8.49	about 11 years
15-19	6.46	20-24	8.93	about 13 years
20-24	7.80	25-29	8.36	
25-29	7.95	30-34	7.54	
30-34	6.80	35-39	6.82	
35-39	5.65	40-44	6.81	
40-44	4.95	45-49	5.64	
45-49	4.35	50-54	4.60	
50-54	4.15	55-59	4.50	
55-59	3.40	60-64	3.48	

norms which would be established from the large standardization sample. Unfortunately the unexpected problems over the difficulty level of the test which resulted in the development and testing of new item material and the eventual changing of the item content of the test made this virtually impossible. Predictions of means of the various subtests would have been relatively simple but predictions of standard deviations, both of subtests and of the three total scores, would have been extremely difficult if it could have been done at all.

The best that could have been done would have been to take a random half of one of the two sections of the standardization sample and compare norms projected from that to norms projected from the other section of the sample. This would have meant developing norms for a set of about 275 examinees, developing another set of norms for a set of about 550 examinees, and comparing these. The other 275 or so cases would have had to be dropped because they would not have been drawn from an independent sampling phase. It was considered that the sample sizes involved were not sufficient to make such a demonstration convincing enough to warrant going through with it.

This does not mean, however, that no evidence at all was obtained concerning the question of whether or not it is feasible to derive relatively accurate norms from such a small sample was obtained. On the contrary quite a bit of positive evidence that it can in fact be done was obtained. This evidence will be reviewed briefly here.

1. It has been demonstrated that a population can be represented very well with samples as small as 250 to 300 persons. A sufficient procedure is to find some device for drawing a random sample of dwelling units. Such a sample can reasonably be but would not have to be stratified on urban-rural residence. Having such a sample of houses a census of them would have to be taken, a sample of persons of the kind desired selected from these, and then the persons so selected contacted and tested. Assuming the sample of houses is selected as in the case of Puerto Rico by taking small clusters of six in urban centers and 12 in rural areas, the amount of travel involved would not necessarily be prohibitive. It would certainly provide wide representativeness. It should also prove much more economical than contacting and testing a large sample.

2. It has been demonstrated that the test materials and instructions function quite well in a second country, in this case Guatemala. It is true that the test administrators in Guatemala found certain phrases and usages a bit strange. It is to be expected that other areas would also find certain of the language usage a bit strange. It is also to be expected, however, that the test will be understood and that its characteristics will be quite similar whereever it is used.

There is one negative element in the claim that the test materials function well in Guatemala. Some of the items in the final form were not tried out there. There is no apparent reason to believe that these newer and generally easier items would fail to serve the same function there that they did in Puerto Rico. Neither is it necessary to believe that they will do so in every respect.

In summary, then, the clear success we have had in applying the test in two Spanish speaking areas should permit reasonable optimism that success with it can be achieved in many such areas.

3. The test can be expected to be more difficult than would be optimum in areas having substantial levels of illiteracy or substantial proportions of persons who have had little opportunity to attend school for more than three years. This fact will render profiles of subtests of doubtful utility. Even so, it has been demonstrated that the Verbal, Performance, and Full Scale scores are clearly adequate to provide reliable differentiation between mental ability levels of people of low educational history.

4. The item statistics for the 376 reference group were in almost all respects highly similar to the item statistics for the full sample of 1127 examinees. This speaks well for the probable stability of item and other test statistics as one shifts from using a restricted age band to the use of a much broader age band. This is important because if a relatively small sample is used, then one would certainly want to use the entire sample both for calculating item statistics and for working out sets of norms to be used.

5. IQ conversions can be established for varying age spans by utilizing the fact that such age level variations have been demonstrated to be due almost entirely to difference in average level of education. By subdividing the sample into about four age level sections a reasonable estimate of a time related regression equation of scores on education can be established. The scaled score value which should correspond to an IQ of 100 can be determined from this regression equation (probably non-linear) for each age level for which norms are desired. It is suggested that this regression be for scores on the average educational level of each age group in order that this might also be utilized to provide some direct information as to the rate of change of scores in the population and therefore as to the time span across which the norms can be expected to be reasonably accurate.

There may have been other evidence that could have been pointed to. It seems clear that the evidence obtained so far is quite favorable and is sufficient to permit people in other areas to expect that such project will succeed.

IV. DISCUSSION

In the main the results of the project have been analyzed and interpreted as they have been presented. This was done in order that succeeding results could be kept in proper perspective as they were presented. The weight of the evidence presented was that the test development project in Puerto Rico has in fact produced a meaningful and useful measure of adult intelligence. Perhaps the most serious questions that can be raised about the quality of the test concerns the fact that various of the subtests do not have an ideal distribution of item difficulties and the fact that the test is still over-all more difficult than would be ideal for the Puerto Rican population. In spite of these two questions, it is clear that all other statistical properties of the test are at least as favorable as were obtained for the English language version originally developed in the United States. It is also clear that all subtests, except Digit-Symbol Substitution, scaled well enough as to permit the three total scores to provide quite acceptable degrees of differentiation between examinees who score low on the total test.

Any difficulty problem which might be present can be expected to be self correcting with the passage of time. The educational level of the population is increasing rapidly. Every subtest correlates substantially with the reported number of years of education of the examinees. This phenomenon is so strong that time related trends in the population can be almost completely accounted for in terms of the formal educational background of the varying age levels.

The self correcting expectation actually implies a progressive invalidation of the norms presented with this report. As the educational level increases, the mean scaled scores earned by a group will increase, and IQ conversions will also increase on average. It is estimated that within about 10 years the 16 to 19 age group will have increased in average education by as much as 2 and 1/2 years. This means that the full scale mean IQ would have become in error by about 6 points. The Scaled Scores will have increased by about 12 points. By the time the mean error becomes this great new norms probably should be substituted for those being published here. The rate of change in the 16 to 19 age group after the 10 year period will probably be due more to changes in the efficiency and quality of the school system than to any further increase in average education. Because of this within 15 to 20 years, a new sampling involving about 250 cases should be carried out to re-establish an accurate set of norms. Other adjustments for other age spans are discussed in section III, E, 2 of the Results.

One can wonder about the optimum manner to arrange for such adjustments. They have not been carried out to the point of actually tabulating them for this report. Such tabulating is quite simple. The publisher of the test material probably should weigh the problem of how best to control this problem. Putting two sets of norms in the manual would be appropriate providing specific data for initiating their use were specified now. It is suggested that it might be better to hold off publishing such revised norms until such time as the Office of Education of Puerto Rico can certify that the younger age groups are achieving approximately 2 more years of education than they now are. While it is possible to be more specific about the changes that will be needed for all people who are now 30 years of age or older, to put in such conversions for some age bands and not for others may prove either awkward or confusing or both. It may be best to simply publish a revised set of norms for all age bands about 1975.

All the statements made about time related changes in scoring and about adjustments in norms to be made in about 10 years derive from the observation that Education, defined as the number of years of school completed, is such a powerful determinant of test performance. Education, defined this way, accounts for 50% of the variance of the scores. Education accounts for virtually 100% of age related trends in scores. Education can be described as having pervasive and permanent effects. It effects every kind of behavior called for on the WAIS. Its effects tend to determine the intellectual level of activity of a person for the rest of his life, or at least to the age of 65. The results obtained from the portion of this project dealing with education balanced age levels, the parent-child data, and from other segments of the data strongly support the idea that the role of education has been considerably underplayed in the thinking of people working in the area of ability testing. They, in turn, have influenced the thinking of other people. Undoubtedly, one of the most important outcomes of this project will be the publication of strong evidence which will go a long way towards changing ideas and beliefs in this area.

A second important outcome of this project is that it had been demonstrated that random sampling methods can be used to obtain samples of perhaps 300 persons who will in fact represent a population very well. This fact coupled with the findings from the Guatemala extension of the project demonstrate that the test will be transportable to other countries on a reasonably economical basis. It is clear that the over-all characteristics of the test (except for difficulty level) will be quite stable from one area to another. It is also clear that "local" norms will have to be established because of variations in scorable responses, language usage, and difficulty levels of individual items, Because of item difficulty variations, it probably will be necessary to adopt less stringent stop instructions for subtests in which these are used.

Projects to develop norms for other areas will almost surely be undertaken. Such efforts probably should be influenced by the fact that almost anywhere it might be done, the test will be on the difficult side. This statement is based on the expectation (which may not be true) that the average level of education is no greater in any Spanish language area than it is in Puerto Rico.

The problem with the difficulty level of the test is presumed to be related to lack of education and therefore to particular kinds of experience and not to the culture or the potential of the people. It is such lack of experience which results in such findings as the fact that some people cannot produce scorable responses to the Vocabulary Subtest. Might avenues have been found which would permit people in effect to learn the desired type of response within the testing situation? A great deal of concern and debate was carried on within and beyond the project group over the instructions and their possible role in the failure of some people to respond. There is no clear evidence that the instructions contributed to such failures but neither can we demonstrate that they did not.

One important element in this regard is that a search for markedly different instructions which might have induced more people to yield more scorable responses would inevitable have meant deviating further from the model that we were using, the format of the U. S. WAIS, that was considered desirable. For better or for worse this model had been adopted and the goal of the project was to produce as good a test patterned after this model as it could within the limitations of time and money provided for it.

It is not claimed that the best possible test for use in Spanish language areas has been developed. There are no necessary grounds on which either to affirm or to deny that the WAIS was an optimum model to follow. Since, however, it is by no means certain that the use of a model that evolved and had been relatively successful in the U. S. culture will also prove to give optimal results in a Spanish culture, it seems desirable that some space be given to a more thorough discussion of the issues involved than has been presented up to here.

Most people who are familiar with the two cultures will readily agree that there are real and important differences between them. There are differences between perspectives and attitudes concerning such important elements as time, space relationships, status symbols, etc. It would surely be more difficult to get agreement on the question of whether or not the existing cultural differences should lead to differences in operational definitions for the measurement of intelligence. If there are pervasive differences in the kinds of behavior that are looked upon as being intelligent, or representing intelligence, then, ideally, intelligence test content should reflect these differences. If differences in the kinds of behavior, or behavioral acts, that are regarded as intelligent are not pervasive, or are relatively limited, then the matter of test content would be of little or not importance.

It does appear that there are cultural differences that could be important to the determination of what the content of the test should be. There are, for instance, differences in attitudes toward education and educational achievement. It is the impression of the present writer that at least in the growing middle class of Puerto Rico these attitudes are changing towards those that prevail in the United States. Be that as it may, it does not follow that attitudinal differences which may exist would necessarily lead to critical differences in what behaviors will be regarded as intelligent. That is, the differences may involve relative emphasis or importance, rather than content. Certainly the importance of the education variable in the Puerto Rican data support rather than negate the use of the WAIS model.

Another avenue by which this question might be approached is to review the definition of intelligence that was implied when the WAIS was chosen as the model to be used. The definition implied was, of course, the one adopted by Wechsler. Actually Wechsler offers various levels of definitions of intelligence from highly theoretical to purely operational. He starts with the theoretical by asserting that intelligence is not a tangible entity but rather is a limiting construct that is known by its effects or properties (2, p. 4). It manifests itself through learning, adapting, reasoning, and other ways. Even though it is manifested in many ways, Wechsler points out, there must be some basic element, or factor, or similarity that is common to all of these ways. It is necessary that this be true if the anxioms of arithmetic are to be used in scoring intelligence tests and deriving intelligence quotients from the process. "...The combining of a variety of tests into a single measure of intelligence, ipso facto, presupposes a certain functional unity between them." (2, p. 10). From the viewpoint of measurement theory his comments so far are quite true. Neither, it would seem, are they culturally bound.

There is, however, a further implication of this position that could be relevant to the culture question. That is that, since all materials deemed to require intelligent behavior exhibit this functional unity, it does not matter which behaviors are required of the subject. All that is required is that the set of behaviors used be sampled from a broad enough portion of the total to give a relatively fixed meaning to the notion of general intelligence. Does the WAIS, as developed in the States, in fact provide a model which is a reasonable sampling of intelligent behavior in Spanish culture? All the statistical evidence concerning this question presented in this report are consistent and strongly support an affirmative answer.

The above theory that all intelligent acts contain *something*, the nature of which is unspecified, in common is really as far as we would have to go in our review of the meaningfulness and appropriateness of the theory. It is a sufficient, even though perhaps not a very satisfying, base on which to accept the WAIS as a model. If the theory is ture, the WAIS is acceptable as a model.

Wechsler, however, attempts to discuss what this *something* is and it probably is worhtwhile to review briefly this part of the theory also. He embraces Spearman's "g" factor as being a *measure* of the mind's capacity to do intellectual work and also as a necessary and important sign of general intelligence. The "g" factor is not, however, considered the only or even the paramount sign. He lists other non-intellectual factors such as social adaptation, perseverance, interest, desire to succeed, and energy level as being of much importance. In other words, intelligence is one aspect of the total personality, an aspect that is related to other factors of personality.

Having said all these things, however, it is still true that the measurement of intelligence is restricted to the measurement of certain mental or intellectual abilities. In fact, Wechsler finally gives an operational definition of intelligence which is essentially the same as Binet's. "Intelligence, operationally defined, is the aggregate or global capacity (ability) of the individual to act purposefully, to think rationally, and to deal effectively with his environment... By measurement of these abilities we ultimately evaluate intelligence." (2, p. 7)

Note that there is certainly nothing that is culturally bound included in this definition. What could be culturally bound would be the sets of purposeful or rational behaviors that are regarded as effective and therefore as intelligent by the members of the culture. Such behaviors are inevitably, either directly or indirectly included in the educational institutions of a national group. The fact that education correlates with test score in Puerto Rico to about the same extent as it does in the States strongly suggests that the WAIS is equally as good as a model in Puerto Rico as it is in the States.

Finally, in considering the question of suitability of the model adopted, we should note that the WAIS has eleven subtests which sample quite a variety of behaviors. These eleven subtests involve certain materials that clearly are directly related to formal education, such as the later items in Information, Vocabulary, and Arithmetic. (As an aside here, we should remind ourselves again and again that all eleven subtests are profoundly affected by formal education). Some subtests are in part measures of immediate memory, such as Digit Span, Digit Symbol Substitution, and Arithmetic. Still others involve verbal skills such as Vocabulary and probably Analogies. So-called Social Intelligence or awareness is represented by Comprehension and Picture Arrangement. Similarities requires an ability to abstract. Abilities to manipulate and organize non-verbal concrete materials are required by Block Design, Object Assembly and Picture Arrangement. One would be hard pressed to show that any of these abilities are not considered representative of intelligence by people of Spanish culture. Are there really crucial gaps in this list? If so, they are not immediately apparent.

Even if there are no crucial gaps in the list, one might still wonder if they are appropriately weighted into the total. The WAIS system of scoring provides that each of the eleven subtests be approximately equally weighted. Cultural differences conceivably might be such that a different system of weights would be optimum. We have made no effort to investigate this possibility.

Finally, there are also differences in certain demographic characteristics of the population which might influence choice and weights of subtests. Certain subtests have been thought to be more closely related to formal education than others, such as Information and Vocabulary, although this idea appears to be false. These subtests have been assigned the same weight in the EIWA as in the WAIS in spite of the fact that there are still substantial differences in average years of education and opportunities for formal education. There are also substantial differences in percentages of urban-rural residents, there being 30% rural in the U. S. and 52% rural in Puerto Rico. It has been claimed at time that tests developed by urban residents are likely to be unfair to rural residents. Such a bias is possible and may be more important in Puerto Rico than in the United States. On this possibility we have little evidence. Rural scores may be lower only because rural residents go to school less.

V. CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

From all that has been presented, it is reasonable to conclude that all four of the general goals that this project set out to achieve have in fact been achieved.

1. There has been produced and standardized in Puerto Rico an individual intelligence scale, written in the Spanish language and adapted to Spanish culture for the age range 16 through 64. The test produced was an adaptation of the Wechsler Adult Intelligence Scale which was first developed in the United States. All materials which could have been considered to be influenced by culture were either eliminated or rewritten in parallel form with Spanish culture as the referent.

The statistical properties of the test are in virtually all respects very similar to the statistical properties of the original English language form. Reliability of the test is excellent. Concurrent validation has been demonstrated. Probable predictive validity in academic settings can be expected because of the extent of correlation of the test scores with number of years of school completed. The one area in which the characteristics of the test might be questioned concerns its difficulty relative to the population in which it is to be used. Several of the subtests were more difficult than desirable. All but one, however, could be scaled to a satisfactory degree and as a consequence the three total scores, (Verbal, Performance, and Full Scale) all could be scaled quite well. There is, then, good differentiation between relatively low scoring individuals. Another difference between the model and the final form of the Puerto Rican test is that no time bonus was assigned for quickness of response in the Arithmetic subtest. This was done because time bonus seemed to have some negative effects on reliability when testing persons in the younger age levels (16-24). Since the test will be used more among such persons than among older persons this kind of outcome was judged to be undesirable. Some time bonuses used for the English language model were also judged to be desirable in other subtests of the Spanish test (Block Design, Picture Arrangement, and Object Assembly). Speed of response, then, contributes to outcome on four of the five Performance subtests and none of the Verbal subtests. Finally, it was demonstrated that four of the subtests could be used as a quite satisfactory approximation to the full scale score if a situation should arise in which this is all that the examiner needs or can get (Vocabulary, Arithmetic, Picture Completion and Picture Arrangement).

2. The second objective of the project--to develop a test and a procedure which will make it possible to adapt the test to other Spanish American countries with a minimum of cost and effort--was also accomplished. The demonstration of this took two forms. First was the testing of a sample of subjects in Guatemala City. From this effort it was clear that the test content, instructions, and most of the scoring keys could in fact be successfully applied in another Spanish speaking country which, of course, had various distinctive elements in its cultural and racial history. In some instances scoring keys needed modification to take local usage into account and there were also variations in item difficulty levels, especially in Vocabulary. In almost all respects the statistical properties of the test were the same in Guatemala City as they were in Puerto Rico.

The second approach to this problem concerned the effort to demonstrate that relatively small sample of from 250 to 300 persons could in fact represent a total population if such samples were obtained by random sampling methods. This effort succeeded very well. Such methods almost surely obtain samples that are more representative of a population that can be obtained by setting up some arbitrary system of matching individuals to census data, especially if the census data is more than two or three years old.

Finally, an outline of procedures was offered by which norms for a population can readily be derived from such a sample rather than by using the present method of adopting some subset of the population to use as a reference group for determining Scale Score equivalents for raw scores.

3. The third objective of the project was satisfactorily accomplished. Information was developed from which automatic corrections can be made in about 10 years (preferably 1975) which will compensate for time related shifts in mean Scaled Scores that are occuring in the population. These time related shifts were demonstrated to be almost wholly accountable for by changes in the average amount of education achieved by each succeeding age group as it reaches maturity. This conclusion was supported with evidence obtained from the standardization sample itself, special education balanced age groups, and parent-child pairs selected and tested for the purpose. All this evidence was consistent and pointed to the conclusion that mean Scale Score for the Full Scale increases just about an even 5 points for each year of increase in the mean education of the groups. Since a change of about two Full Scaled Scores is equivalent to a change of one IQ point, and since age groups between 16 and 29 can be expected to have increased their average level of education by at least two full years by 1975, the norms will be off by an average of at least 5 points by then. Since major efforts are also being made to improve the schools

of Puerto Rico, it may well be that such changes will even be larger over the next decade. For this reason, the norms should be checked about 1980 by obtaining a new sampling of about 250 persons. This sample should be representative of the entire population in order that age level norms can be checked at all levels.

4. The fourth objective--to incorporate into the overall effort the investigation of a variety of questions of general interest to Psychology--could have been stated more broadly and included Education. This objective has been achieved in part and will be more fully realized in the months that follow. It is informative in this regard to outline some of the substantive results that have already been brough out.

a. Probably the most important single finding so far is the fact that age related trends in Full Scale score from age 25 to 64 can be accounted for almost completely in terms of the educational levels of the age intervals being compared. To the extent that there are age related trends, they consist of a tendency for Verbal Total to rise until about age 39 and then never to fall even to beyond age 64! Performance Total, on the other hand, shows no statistically reliable trend either way from age 25 through age 64 when the education variable is controlled. This is of importance because of widespread belief that the scores one earns on intelligence scales are substantially uneffected by the educational background of individuals especially after they have been out of school for quite some time. The usual argument is that mental ability peaks in some sense at about age 30 and then declines slowly thereafter. The decline is argued to be faster in some functions than others. These results negate these beliefs. Mental ability clearly does not decline to any measurable extent until well into the mid-sixties. When such decline may actually begin is not predictable from these data.

b. Subsidiary to the above finding is the fact that in education balanced groups there is no evidence that immediate memory functions as represented by Digit Span, Arithmetic, and Digit Symbol Substitution decline before about age 65. In fact, in the case of Arithmetic this was not even evident in age groups which were representative of their educational levels in the population. Clearly, this concept of declining immediate memory will have to be reexamined. There is no immediate or even remotely available explanation as to why such a function should exist in one culture and fail to appear in another. Since a declining function does not appear in the Puerto Rican culture one has to assume that such a function in our culture is an artifact of the culture and cannot be taken to represent any necessary physiological diminishing of potential to maintain such an ability. The fact that it is an artifact in this country is supported by the fact that Digit-Symbol Substitution, which is a closely timed speed test, as well as involving immediate memory functions, does show age related decline, even in education balanced groups. It would appear that the tendency to slowing of physical reactions and speed of movements does in fact represent physiological changes that occur accross cultures.

c. The demonstration of the power of the random sampling methods used to produce fully representative samples of a population should be of considerable interest to researchers in Psychology, Education, and related disciplines. The same methods or various variants of them can be adapted for use in other settings even though the interest of the researcher may be much more restricted in scope than the entire population of a country.

d. The results of this project offer one more bit of support to the idea that intelligence can be defined in essentially the same way as one goes from one culture to another. There are those who argue that there are perceptual, experiential, attitudinal, interest, and other differences that must inevitably exist between cultures that require that new test development projects should be based on definitions which are developed strictly in terms of the culture in which the test is to be applied and that projects such as this one are inappropritate. Of course, to date the only clear validity evidence obtained has been the concurrent validation phase of this project. As the predictive validity of the test is investigated in a variety of situations these arguments will be brought into sharper focus.

e. At least 10 investigations of various problems which will be of general interest both in Psychology and in Education will be carried out on various portions of the data that was accumulated during the course of this project. Some of this work is already under way.

5. Before closing this summary one final matter should be briefly discussed. During the course of this project a persistent worry, esepcially of the Director, has been that it might be that a test like this applied in an area in which opportunity for education is still quite unevenly distributed through the population might serve to perpetuate the relatively low state of many people who could, if given appropriate opportunity even during early adult years, develop and achieve at much higher levels. It seems highly desirable that investigations should get under way which would be directed to the possibility of using test information in ways which would avoid the possibility that this information might have inappropriately negative outcomes for the individual with limited

opportunity for education. One such avenue that can be tried is to investigate the predictive validity of scores that have been regressed on education, especially among low education people. Since fully 50% of true variance has not been accounted for by the education variable, even in the total population, there is more than ample leeway for such regressed scores to prove effective far beyond the pessimistic predictions that would be made from the unregressed scores. It is hoped that such a project can be set up in Puerto Rico in the near future. Such a project would be a natural sequel to this one. The idea involved have grown out of considerations which arose during the labors of this one.

VI. SUMMARY

1. The problem

So far as could be determined, no properly developed and standardized adult intelligence scale had been developed in all of Latin America. It has been common, as a result of this for those who needed such an instrument to translate a scale from another source and then to apply norms that were inappropriate. For this reason the Faculty of the Department of Psychology decided to undertake a project to adapt and standardize the Wechsler Adult Intelligence Scale in Puerto Rico.

2. Objectives

There were four general objectives which it was anticipated that the project would achieve.

a. To produce and standardize in Puerto Rico an individual intelligence scale written in the Spanish language and adapted to Spanish culture for the age range 16 to 64.

b. To develop a test and a procedure which would make it possible to adapt the test to other Spanish American countries with a minimum of cost and effort.

c. To develop information which would allow setting up automatic corrections which would compensate for the time related shifts in mean score that is normal for tests of this kind.

d. To incorporate into the over-all effort the investigation of a variety of questions of general interest to psychology.

3. Methods

By permission of the Psychological Corporation, the project undertook the task of adapting the Wechsler Adult Intelligence Scale (WAIS) for use among Spanish speaking adults. The advantages of starting from a widely known and used model for such a project are many, not the least of which is the fact that it makes it possible to avoid repeating large amounts of work in definition, formulating guides for scoring, and development of items material. One of the most important ingredients for success was present. The Psychological Corporation agreed that item content could be changed as much as required to achieve a true adaptation to Spanish culture.

a. Translation and adaptation

First the test was translated directly into Spanish. The translation was edited twice. The edited translation was then gone over carefully, and all material that could in any way be challenged as being biased toward the culture of the United States or as being biased in that it would be unfair to the culture of Puerto Rico or of Spanish speaking people in general was either eliminated or changed. Some such changes simply involved using alternate forms having Spanish culture as the referent. At all time the goal of using Spanish vocabulary, syntax, and materials which would be understood and applicable throughout the Spanish speaking areas of the hemisphere was kept before the group. Local usages were carefully avoided.

b. Assembly and review of the preliminary manual

The revised and new materials were assembled into an experimental manual which was about 25% longer than the final test was expected to be. Two linguists specializing in the study of local and regional usages of the Spanish language reviewed the materials for correctness, simplicity, intelligibility, and generality. Psychologists who were active in test development in Guatemala, Venezuela, and Uruguay were then asked to review the manual using the same criteria. The suggestions and criticisms of all were collated, and in a few instances compromised between conflicting opinions. It is clear that the manual was in fact easily understandable in areas far beyond Puerto Rico.

While the above reviews were still in process, preliminary testing of test materials was started. Written responses to the Vocabulary words were obtained and used for a preliminary screening of words and for preliminary setting of scoring criteria.

c. Training and pretesting

Examiners were recruited, intensively trained, and individual testing was also begun. Most examiners were Masters Degrees level Psychologists. The initial individual testing following training was of two widely disparate groups which were composed of 24 University students and 20 individuals who had been diagnosed by a Ph.D. Psychologist as being of borderline normal to high level defective in intelligence. This data was made to make an initial course screening of item discrimination and to give an estimate of the difficulty level of the test materials. It was expected that the mean of the population should fall more or less half way between the means of these two groups.

d. The "254" or item analysis sample

The information gained from the reviews and the preliminary testing was combined to organize a manual which was to be used for a definitive effort at item analysis and investigating in full the characteristics of the test as developed for the Spanish culture. This involved obtaining a representative sample of the island population which would consist of about 250 persons.

The 254 sample was obtained by using a multi-stage, partially urban-rural controlled, random sampling method. It was demonstrated that the method did in fact represent the population very well.

Each person in the sample was contacted and his cooperation solicited. The refusal rate was only 5%, although a total loss for reasons such as mental illness, severe physical illness, having moved to the United States, etc. was about 20%. This 20% was replaced with similar persons who had been drawn in a second back-up sample obtained specifically for this purpose.

The data from these 254 testings were used to obtain estimates of item difficulty, compute, item-test correlations, investigate the homogeneity of item content, investigate the relative validity of alternate scoring keys for various items, investigate the merits of certain of the time bonuses used in the scoring keys, and to compare the Puerto Rican test performance characteristics with those found in the United States.

The critical outcome of the 254 sample analysis was the discovery that the difficulty level of the test was too high. The population average turned out to be at the level of the low group in the disparate group testing.

e. The Guatemala extension of the project.

While the 254 testing phase was in process arrangements were also made for a sample, also randomly drawn, to be tested in Guatemala City, Guatemala. This data was used to obtain an initial assessment of the generality of the test instructions and item materials.

f. The "124" or second item analysis sample.

Since the 254 sample had shown that the scale was too difficult and after consultation with the Psychological Corporation and the Office of Education, it was decided to initiate a new and unexpected cycle of item development and analysis. New, easier material was developed and some difficult material was dropped. A new random sample, restricted to the northeast fifth of the island, was drawn and tested, and a new round of analyses completed.

g. The standardization

The results from the preceding phase having shown that the test now would function satisfactorily, the project proceeded to the large standardization effort. A standardization sample of 1176 persons was drawn by the random methods used previously. Actually, the sample was drawn in two sections, the first early in 1965 and the second in early summer of 1965. The individuals were again contacted and 1127 testings were completed. Refusal and replacement rates were almost exactly the same as before. The standardization sample also proved to be quite representative of the population of the island.

When a completed protocol was delivered to the project offices, its scoring was checked for accuracy, all demographic data, all item scores, and all 11 subtests and the three total scores were transferred to IBM cards. At least two 100% accuracy checks were carried out on all phases of data handling. Errors were, it is believed, reduced to an extremely low level. Data analyses were carried out on an IBM 7074 computing facility under the direct supervision of the Director of the Project.

4. Results

a. The "254" item analysis data

All analyses were very favorable except for the difficulty level of the test. Item versus subtest total correlations (corrected for contamination) were generally quite high and quite satisfactory. Reliability of the subtests was high and at least as favorable as had been obtained on the English language model. Correlations among the subtests and the three total scores were all similar to those found in the States.

Because of the difficulty problem several analyses involving subdivisions of the population which might yield more favorable results were carried out. These involved age, education, and urban-rural residence in particular. No subdivision tried prove sufficient to overcome the difficulty problem. However, these analyses did show that sufficient items functioned properly among low education people that we could expect that meaningful differentiations among them would be made. They also showed urban-rural differences that persisted even after number of years of education was accounted for. They also showed sex differences which appeared to persist beyond the fact of somo difference in education between the sexes. Finally, and very important to note, was the finding that some 50% of the Full Scale score variance could be accounted for in terms of number of years in school and that every subtest correlates substantially with education

-- the Performance subtest doing so somewhat more than the Verbal (except for Information).

b. The Guatemala extension

The Guatemala sample differed from the Puerto Rican in that it had an average of 5 years of education as compared to 8.5 in the Urban zone of Puerto Rico. The Guatemalans averaged two years younger also.

The Guatemala data was subjected to almost every analysis carried out on the Puerto Rican (254 sample) data. Item-analysis, subtest intercorrelations, and reliability results were all very similar to Puerto Rican results. There were differences in apparent difficulty level of whole subtests and of individual items. The Guatemalans actually scored as well as the Puerto Ricans on Vocabulary, Arithmetic, Block Design, Picture Arrangement, and Object Assembly. Guatemalans with 4 years or more of education scored higher than comparable Puerto Ricans on all subtests except Comprehension. In fact for people of 4 or more years of education the difficulty level of the test was almost appropriate. The presence of a higher percentage of illiterates, however, would require that the test be adapted for the whole range of educational background.

c. The "124" items analysis sample data.

New test material was developed, a new manual prepared, a new sample drawn and tested, and the data subjected to analyses similar to those of the 254 sample. Difficulty level of the test was improved but still not ideal. However, since it was apparent that a workable and useful test had in fact been achieved, it was decided to proceed to the major standardization effort. In retrospect, if the test had been made much easier it seems likely that it might have become too easy to be ideal for Guatemalans who have had more than 4 years of education.

d. The standardization sample data

Means and standard deviations of age spans 16-17, 18-19, and each 5 year interval thereafter were calculated. It was apparent that age related trends were somewhat different from those observed in the United States. Verbal scores tended to rise (or at least show no reliable trend) until later ages and Performance scores were highest at age 17 and declined rapidly thereafter. After discussions with the Psychological Corporation it was decided to adopt the same 20-34 age span used in the United States and for similar projects in

Europe and elsewhere as the reference group for determining Scale Score equivalents for the raw scores earned on the various subtests.

A new investigation of the effects of time bonus scores on the item statistics of the Arithmetic subtest was carried out. This investigation substantiated the findings observed in the 254 data that such bonuses tend to lower slightly the reliability of this test among young people. The reliability of older people was essentially unaffected. For this reason time bonuses were omitted from Arithmetic.

Item analyses demonstrated that all items were constributing satisfactorily to each subtest total. All items were, that is, except for a few extremely easy items for which automatic credit could be given if certain later items were responded to correctly.

Scaled Scores equivalents for all raw scores of all subjects on all subtests were then obtained. These were subjected to various analyses. All such analyses proved very favorable. Reliabilities tended to be slightly higher (Full Scale = .98) than those observed in the United States (.97). Intercorrelations among subtests and the three total scores were all similarly high.

Differences that occurred between the means and standard deviations of the reference group and the full sample on each of the 6 Verbal subtests are trivial. The means of all except Block Design of the 5 Performance subtests are reliably lower in the full sample than they are in the 20-34 age group. In both cases standard deviations are quite similar. It is clear that using the entire sample as the reference group for determining scale score equivalents of raw scores would have yielded a somewhat different but no less favorable outcome.

IQ equivalents of the Verbal, Performance, and Full Scale totals were obtained by using a simple linear conversion to distributions having means of 100 and standard deviations of 15.

Concurrent validity of the test was investigated in essence, by standardizing in Puerto Rico a Spanish language form of the Otis Higher Examination, Form A, which had been prepared in Guatemala. The test was administered to every person in the second section of the standardization sample who had had 7 years or more of education. There were 205 such individuals. All but two of the 75 items in the Otis proved to correlate significantly with the total score on the test minus that item. Sixty-two of the items correlated high enough to be considered satisfactory. The Otis Score correlated .856 with the WAIS Verbal Score, .725 with the WAIS Performance Score, and .853 with the WAIS Full Scale Score. Considering the restriction on education involved, these correlations are excellent.

Finally, the question of which set of the subtests might prove in some sense to make an ideal short form for the WAIS was investigated. It was shown that Vocabulary, Picture Arrangement, Arithmetic, and Picture Completion predict the Full Scale true variance better than any other set of four (R = .976).

e. The need for automatic adjustments of norms.

An effort to predict the rate of change of IQ scores in the population was made. To do this two additional kinds of data were obtained. The first consisted of drawing extra cases at the various age levels in order to set up four educationally balanced age groups. The four age groups (25-29, 35-39, 45-49, and 55-64) averaged 7.7 years of education, with the oldest group being slightly lower than the other three. This was close to the average in the total population.

The means and standard deviations of the 14 variables were obtained. The Verbal Total increases significantly with age (F = 10.42, d.f. = 3.536)! This is contrary to expectations. The Performance Total does not show a statistically reliable time related trend. This too is contrary to prior expectations which were that they should be declining. The Full Scale Means of the Scaled Scores do not show a statistically reliable time related trend!

It would appear that by far the dominant influence in age related trends is the Education Variable, at least for the ages from 25 to 64. This being the case the next step can reasonably be to attempt to estimate to what extent scores vary with education. This problem was approached in three ways. The first was to obtain the exact mean education of each succeeding 5 year age group in the standardization sample. The 16 to 19 group was divided into two groups. Across the ages 16-30 there is neither an apparent trend in the number of years of education nor is there an apparent trend in the Full Scale Scores of the groups. From age 30 on to 64 Full Scale Scores of the groups decline, the decline being due to decline in Performance total. The change across this span, when stated in terms of average education of the groups rather than their ages, appears to be about 5 Scale Score points per year of education.

A second approach to the problem was to sort the standardization sample examinees into groups according to educational levels, allowing age to vary in the groups. For this analysis, however, the age span 55 to 64 was omitted. This analysis presented a clear picture of Verbal Scores rising linearly with education at the rate of 3 points per year, of Performance Scores rising linearly with education at the rate of 2 points per year, and, therefore, of Full Scale Score rising linearly with education at the rate of 5 points per year.

The third approach was to utilize a second set of special data obtained during the testing. One of the parents of a random set of the examinees of ages 16 to 25 was tested. Similarly among a random set of parents who appeared in the standardization, one of any children living at home and aged between 16 and 25 was tested. In either case the selection of the person to be tested was random. The means and standard deviations of three age divisions of these parent-child groups were compared. Groupings were based on parents' ages and were by 10 year intervals starting at age 35. The average ages of the children of these three groups varied by only 3.4 years. Again no age related differences were found. Education related differences were found and were 3.56 Full Scale point for each year of education difference.

It would appear that the parent-child differences might be underestimating the point change related to education. The parent-child group means are, of course, correlated. Parents and children live in more homogenous settings than would be true of independent groups obtained at each age level. Hence it seems more likely that the education related Scaled Score changes are closer to the estimate of 5 points than to the 3.56 points.

From census data and from the age trends apparent in the standardization sample it is clear that the rate of change of average education has been about 3/4 of a year for each succeeding 5 year age span for several decades. It also appears that there has been some accelaration of this rate recently. Consequently one can expect that within 10 years (1975) the Full Scale Scales Score mean of younger age groups will be off by about 10 points and therefore the IQ's will be off by at least 5 points on average. New norms should be published when this occurs. Since there is also a concerted effort going on to improve the schools this estimate may be too low. For this reason, it is recommended that about 1980 a new sample of about 250 persons should be selected and tested to make a more precise adjustment of the norms.

f. The prediction of norms from relatively small samples.

The feasibility of this has been demonstrated from several aspects of this study. The sampling methods produce representative samples. The item and test statistics are similar no matter whether the reference group is aged 20-34 or is the full sample. Since age related changes are accountable for almost entirely from the education variable it is relatively simple to estimate the means of various age bands from fitted curves. Finally, the success in Guatemala makes it clear that the test can in fact be successfully applied in areas other than Puerto Rico.

5. Conclusions and Implications of the results of the project

a. The project has in fact successfully produced an adult intelligence scale in the Spanish language and adapted to Spanish culture which is appropriate for use from age 16 through 64. It can be criticized to some degree on the grounds that it is somewhat too difficult but this defect will be self-correcting as the educational level of the succeeding age groups increases and as the quality of the schools improves. Both of these changes are in fact occuring.

The statistical properties of the Spanish language WAIS are quite similar in all other respects to those of the English language WAIS. In some cases they are even slightly more favorable.

b. The test has been demonstrated to be transportable to other Spanish speaking areas. The sampling methods used show that this can be done reasonably economically of time and money.

c. It was demonstrated that there would be need to adjust the norms about 1975, that this adjustment would be about 5 IQ points at the younger age levels, and that for older groups it can depend on the age of the person at the time the test was standardized (1965).

d. Several questions of general interest to Psychology and to Education were investigated with some surprising results.

Perhaps the most startling result is the strong evidence developed that there probably are no measurable decline in mental abilities, per se, up to age 60 to 64 or so. All age related trends in the data can be accounted for in terms of changes in average education of the age levels. To the extent that there are age related trends they consist of a tendency for Verbal Total to rise until about age 39 and then to remain almost stable at least to age 64. Performance Total, on the other hand, shows no time related trend from age 25 through age 64. The Digit-Symbol Substitution subtest does decline with age but this is primarily a perceptual speed task and the decline is due to slowing of response and movements rather than to decline of mental abilities.

There also is no evidence in these data that immediate memory functions, as represented by Digit Span, Arithmetic, and Digit Symbol Substitution, decline before about age 65. In fact, in the case of Arithmetic the tendency was for the means of the age groups to rise slightly all the way up to about age 60 (although this was not a statistically reliable trend) even when education was representative in the groups rather than balanced. Since such a decline does not occur in Puerto Rico, one has to conclude that to the extent that it does occur in the United States it must be due to some artifact in the culture rather than representing any necessary change due to physiological deterioration.

The results of the study offer another set of evidence that intelligence can be defined in essentially the same way as one goes from one culture to another. Item content, of course, must vary with the culture. The factors or abilities or constructs are what appear to be relatively constant across cultures. This is not to say that changes in the definitions of intelligence might not produce a still more appropriate test than the one produced here. There is nothing in these results, however, which would lead one to expect that such would in fact be the case.

Several further investigations of the data collected for this project are being undertaken but since they were not directly relevant to the project itself, they have had to wait until the completion of the project to be undertaken.

One project that really should be undertaken that is a natural sequence to this one is to investigate the possibility that a test such as this one which has been developed in an area in which the opportunity for education is still unevenly distributed might unfairly discriminate against people of relatively little educational opportunity. This can be approached by investigating the variance in age regressed scores to see if they might not lead to more realistic predictions for such persons than the present method of scoring might lead to. It is hoped that such a project will get under way.

BIBLIOGRAPHY

1. Green, Russel F. Desarrollo y Estandarización de una Escala Individual de Inteligencia para adultos en Español. Revista Mexicana de Psicología, Vol. I, 1964, 231-244.

2. ______________. How Meaningful and Appropriate is the Project? Paper read at the 11th Convention of the Puerto Rican Psychological Association on October 8, 1965.

3. Wechsler, David. The Measurement and Appraisal of Adult Intelligence. 4th Edition, 1958, The William & Wilkins Co.

APPENDIX I

NOTICE

The Psychological Corporation, 304 East 45th Street, New York, New York 10017, is the publisher and owner of the copyright of the Wechsler Adult Intelligence Scale (WAIS) which was translated into Spanish and standardized in Puerto Rico. The Manual for the translated version of this test and many of the test items and answer forms appear in this report which is a copy of the project report prepared for the Commissioner of Education of the United States Department of Health, Education, and Welfare in accordance with the provisions of the contract. When the Psychological Corporation consented to this project the University of Puerto Rico and the United States Department of Health, Education and Welfare agreed that the Psychological Corporation would have the right to copyright and publish the Spanish edition and that the inclusion of any test materials in this report would not adversely affect the rights of The Psychological Corporation which reserved all its rights to the basic materials.

These materials are published herein with the consent of The Psychological Corporation. The Wechsler Adult Intelligence Scale is Copyright © 1955, 1963 by The Psychological Corporation, New York, New York 10017.

MANUAL

EIWA

ESCALA DE INTELIGENCIA
WECHSLER PARA ADULTOS

DAVID WECHSLER

Traducción, adaptación y estandarización hecha
en Puerto Rico con permiso de "The Psychological
Corporation", Nueva York, N. Y.

MANUAL

para la

ESCALA DE INTELIGENCIA WECHSLER PARA ADULTOS

DAVID WECHSLER

Este proyecto fue patrocinado por la Oficina de Educación, Salud y Bienestar de los Estados Unidos y el Departamento de Psicología, Colegio de Ciencias Sociales, Universidad de Puerto Rico

PREFACIO

La necesidad de llevar a cabo un proyecto de investigación con el propósito de proveer una medida de la inteligencia, válida para la población adulta puertorriqueña, era evidente en Puerto Rico. La misma necesidad existe en otros países de habla hispana. Este hecho se tomó en consideración al desarrollar la prueba de manera que se facilite la realización de proyectos análogos en esos países. La realidad existente demostraba que no había un instrumento en español con normas adecuadas que se pudiera usar con confianza para evaluar la inteligencia de los adultos. Con anterioridad a esta investigación se hicieron varios intentos para tratar de llenar esta necesidad, a saber: el Centro de Orientación de la Universidad de Puerto Rico preparó una traducción experimental de la Escala de Inteligencia Wechsler Bellevue, Forma I; la División de Neurosiquiatría de la Administración de Veteranos de Puerto Rico también tradujo y adaptó en parte la misma prueba; la División de Evaluación y Exámenes del Departamento de Instrucción llevó a cabo un proyecto de traducción y adaptación de la Escala de Inteligencia Stanford Binet, Forma L, 1937, y de la Escala de Inteligencia Wechsler para Niños. Los psicólogos se han visto obligados a usar traducciones de exámenes de inteligencia sin la debida estandarización en nuestra cultura. Esta práctica podría, en algunos casos, resultar en errores de clasificación de individuos y las consecuencias podrían ser profundas y perdurables.

El número de agencias, instituciones y empresas públicas o privadas que proveen y utilizan servicios psicológicos es cada vez mayor. Por consiguiente es muy necesario tener instrumentos psicológicos de evaluación que sean apropiados y redunden en mayor efectividad de los servicios.

Las razones anteriormente expuestas son algunas de las más importantes entre las que indujeron al Departamento de Psicología del Colegio de Ciencias Sociales de la Universidad de Puerto Rico a llevar a cabo un proyecto tan necesario y retador: preparar la Escala de Inteligencia Wechsler para Adultos (EIWA) la cual presentamos aquí como un instrumento para usarse con la población adulta puertorriqueña. Concurrentemente se ha llevado a cabo un proyecto similar en la Universidad de San Carlos, de Guatemala, patrocinado por la Universidad de Puerto Rico con fondos de la "Cooperative Research Grants".

La realización efectiva de este proyecto fue posible mediante los esfuerzos unidos de la Oficina de Educación, Salud y Bienestar de los Estados Unidos y la Universidad de Puerto Rico con la cooperación de la "Psychological Corporation".

La estandarización de la Prueba de Inteligencia Wechsler para Adultos, (llamada EIWA) se basa en una muestra al azar de 1,127 sujetos, representativas de la población adulta de Puerto Rico en 1965. La exactitud del método de muestreo se cotejó comparando la muestra obtenida con el Censo de Puerto Rico de 1960. La administración de la prueba estuvo a cargo de 16 examinadores, quienes recibieron adiestramiento especial al respecto. A cada examinador se le proveía el nombre, dirección, sexo, edad, instrucción, zona de residencia y ocupación de los sujetos que constituían la muestra. En general, alrededor de 75 profesionales, entre estos psicólogos, profesores, estudiantes graduados y otros miembros del personal de la Universidad de Puerto Rico participaron en el proyecto en alguna forma.

El Manual que se presenta a continuación contiene una exposición acerca de las bases sobre las cuales se desarrolló una prueba como la EIWA, una explicación del porqué del uso de la WAIS como modelo, un análisis de los procedimientos de muestreo que se usaron, una discusión de los hallazgos más sobresalientes y un número de tablas que presentan diferentes tipos de datos estadísticos. Se incluyen, además, otros aspectos indispensables, tales como las instrucciones completas acerca de la administración y valoración de la prueba y las normas para la población puertorriqueña, por grupos de edad. No se derivaron normas para las personas de 65 años en adelante. Se incluye toda la información que el examinador necesitará para hacer una evaluación adecuada de los resultados. Se espera, además, que los investigadores que así lo deseen, puedan obtener suficiente información para llevar a cabo estudios comparativos o suplementarios a esta investigación. El manual no incluye la información contenida en el libro "The Measurement and Appraisal of Adult Intelligence"[1], acerca de la naturaleza de la inteligencia, la significación de las subpruebas por separado y las aplicaciones diagnósticas y clínicas. Los psicólogos que hayan de usar este instrumento deben estar familiarizados con la información que contiene el libro y tener acceso a la mencionada obra. Los directores de este proyecto tienen planes futuros que incluyen el publicar más información acerca de esta versión en español de la Wechsler Adult Intelligence Scale (WAIS).

La realización de este proyecto ha sido una experiencia rica y fructífera para los que hemos tenido la oportunidad de vivirla. Esperamos que esta investigación sea el comienzo de una serie que venga a llenar las necesidades del país en este campo.

---

[1] Wechsler, David, The Measurement and Appraisal of Adult Intelligence. (4th edition, 1958, The William and Wilkins Co.

Expresamos nuestra apreciación a todas las personas que intervinieron en la realización del proyecto ya fuese trabajando directamente en él, mediante labor de consulta, o dándonos su apoyo y estímulo de cualesquier otro medio. Hacemos mención del Dr. Harold Seashore y el Dr. Jerome E. Doppelt, de la División de Pruebas de la "Psychological Corporation"; de los doctores Alice Y. Scates y Richard B. Otte, del Departamento de Salud, Educación y Bienestar de los Estados Unidos; del Dr. Miguel Valencia, Profesor de Estadísticas del Colegio de Administración Comercial, quién trabajó en la selección de la muestra y proporciónó labor de consulta y del Sr. Carlos Marichal, Superintendente de la Imprenta del Departamento de Instrucción, por los diseños hechos para las subpruebas de Manipulación.

Nuestro agradecimiento especial a la Srta. Laura L. Herrans, a la Sra. Elia H. de Christensen, Profesoras del Departamento de Psicología, quiénes desempeñaron la función de subdirectoras en diferentes etapas del Proyecto. Se hace reconocimiento especial a la labor realizada por la señora Fredesvinda Miranda de De Jesús, supervisora del proceso de administración de la prueba primeramente y, subdirectora del proyecto durante las últimas etapas. Nuestro reconocimiento sincero a los señores Gregorio López Machado y Edwin B. Fernández, investigadores asociados que llevaron la mayor responsabilidad durante el proceso de administración de los exámenes. Damos las gracias a las señoras Jane Podstowski que trabajó en la programación y en el análisis de datos en las etapas iniciales y a Nora Rivera de Vázquez, investigadora auxiliar, quién trabajó con esmero en las últimas fases del Proyecto. Reconocemos el gran esfuerzo de todos los examinadores cuya labor tesonera contribuyó a que el proceso de administración de exámenes fuese un éxito.

Agradecemos trambién la cooperación del Departamento de Instrucción, del Departamento del Trabajo, de la Facultad del Departamento de Psicología del Colegio de Ciencias Sociales y de todos los estudiantes universitarios que participaron en esta investigación. Deseamos también mencionar la labor realizada por las señoras Angela Von Gundlach y Rosario Torres Campos, auxiliares administrativos en diferentes etapas del Proyecto.

RUSSEL F. GREEN

JUAN N. MARTINEZ

28 de marzo de 1967

TABLA DE CONTENIDO

Página

Prefacio iii

Lista de Tablas viii

Capítulo I. Introducción 1

Razones que fundamentaron el desarrollo de la Escala de Inteligencia Wechsler para Adultos (EIWA) — El concepto de Cociente de Inteligencia — La desviación del C I — Organización de la prueba.

Capítulo 2. Estandarización de la Escala 7

La muestra de estandarización — Descripción y naturaleza representativa de la muestra — Confiabilidad — Intercorrelación de las subpruebas — Derivación de las puntuaciones a escalas — Derivación de las tablas de los Cocientes de Inteligencia — Interpretación de los Cocientes de Inteligencia — Correlación de la EIWA con la Prueba Otis.

Capítulo 3. Consideraciones Generales para la Administración de la Escala 30

Instrucciones y procedimientos — La labor oficinesca del examinador — Conversión de las puntuaciones brutas a puntuaciones a escala — Método para determinar los Cocientes de Inteligencia.

Instrucciones para la Administración de la Prueba

Información 39
Comprensión 43
Aritmética 45
Analogías 48
Repetición de dígitos 50
Vocabulario 54
Dígito - Símbolo 56
Dibujos para completar 58

Página

Diseños con cubos 60
Ordenamiento de dibujos 64
Composición de objetos 67

Criterios generales para la valoración y ejemplos de respuestas

Comprensión 71
Analogías 81
Vocabulario 86

Tabla de Puntuaciones Brutas y sus Equivalentes en Puntuaciones a Escala 103

Tabla de Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia 104

LISTA DE TABLAS

Tabla Página

1 Muestra de Estandarización por Edad y Sexo 9

2 Muestra de Estandarización: Región por Zona de Residencia 10

3 Muestra de Estandarización por Edad, Sexo y Zona 11

4 A Muestra de Estandarización por Sexo, Edad y Ocupación: Hombres 13

4 B Muestra de Estandarización: por Sexo, Edad y Ocupación: Mujeres 14

5 Muestra de Estandarización por Edad, Sexo y Ocupación 15-17

6 Coeficientes de Confiabilidad y Errores Estándar de Medición (EEM) de las Subpruebas y Escalas Totales 20

7 Intercorrelaciones de las Subpruebas -Edades 16 - 19 22

8 Intercorrelaciones de las Subpruebas Edades 25 - 34 23

9 Intercorrelaciones de la Subpruebas Edades 45-54 24

10 Promedios Aritméticos y Desviaciones Estándar de las Sumas de las Puntuaciones a Escala: Verbal, Manipulación y Total, por Grupos de Edad 26

11 Relación de los C I de la EIWA con Rangos Percentiles 27

12 Clasificaciones de Inteligencia 27

Tabla Página

13 Edad y Nivel de Instrucción de la Muestra Otis..... 28

14 Coeficientes de Correlación entre la Prueba Otis y la EIWA 29

15 Puntuaciones Brutas y sus Equivalentes en Puntuaciones a Escala 103

16 Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia por Grupos de Edad 104

CAPITULO I

INTRODUCCION

Razones que fundamentaron el desarrollo de la Escala de Inteligencia Wechsler para Adultos (EIWA).

Las razones que sirvieron de base para la preparación de la EIWA son obvias. Como ya hemos indicado anteriormente la necesidad de tener un instrumento psicológico adecuado para evaluar la inteligencia de los adultos en Puerto Rico era imperiosa. Lo evidencia el hecho de que por muchos años se han usado pruebas de inteligencia que son inapropiadas para esta población. En muchas instancias los psicólogos y las agencias que interesan instrumentos de esta naturaleza se han visto obligados a utilizar traducciones sin la debida adaptación y normalización. Hasta donde se ha podido determinar, éste es el primer intento formal de traducir, adaptar y normalizar una escala de inteligencia adecuada para la población adulta puertorriqueña (16-64 años) y que al mismo tiempo tenga posibilidades de ser adaptada y normalizada en otros países cuya lengua vernácula sea el español, especialmente los de la América Latina.

Se podría pensar que el uso de normas derivadas en otras poblaciones no es de por sí perjudicial a la población cuya inteligencia se desea medir. Se ha dicho que las traducciones que existen de pruebas de inteligencia para las que no se han derivado normas locales pueden tener el mismo grado de eficiencia predictiva que las escalas originales. Sin embargo, hasta la fecha no existe evidencia satisfactoria que corrobore esta hipótesis. Por lo tanto, es más importante considerar que al usar normas inadecuadas se incurre en la clasificación errónea de individuos y, como resultado, se desarrollan actitudes incorrectas hacía estos. Tales clasificaciones y actitudes podrían tener efectos profundos y perdurables en la vida de esas personas. Estas clasificaciones erróneas y sus efectos posteriores pueden ser el resultado de una validez espuria. Ello se debe, mayormente, a que la medida que se usó no era justa y los resultados del examen revelan no lo que el sujeto pudo haber hecho sino lo que se le indujo a hacer.

Estas razones, entre otras, sirvieron de base para desarrollar una prueba de inteligencia debidamente adaptada y, sobre todo, normalizada en la población adulta puertorriqueña.

La decisión de adaptar como modelo un instrumento que previamente se preparó y se usó en los Estados Unidos en vez de desarrollar una prueba original fue objeto de análisis detenido. Inicialmente se consideró la idea de preparar una prueba completamente nueva. El análisis hecho reveló que las dificultades que surgen en la realización de un proyecto de tal magnitud y naturaleza sobrepasan las ventajas que de él puedan derivarse. Así lo revela la experiencia obtenida a base de la preparación de otras pruebas. En términos de economía, tiempo, practicabilidad y otros factores, este procedimiento no parecía ser el más conveniente. Debido a esto se decidió traducir, adaptar y estandarizar un instrumento ya existente, de reconocida utilidad, en vez de dedicarnos a preparar una nueva prueba.

Hubo varias razones para seleccionar "The Wechsler Adult Intelligence Scale" (WAIS) como el modelo sobre el cual elaborar la Escala de Inteligencia Wechsler para Adultos (EIWA) a saber: los psicólogos de la localidad se deciden por esta prueba en preferencia a otras con propósito de evaluación. Es un instrumento que ha probado ser válido y confiable. La WAIS ha demostrado ser de gran utilidad cuando se trata de problemas de orientación, diagnóstico y pronóstico. Es una política segura y económica seleccionar un instrumento que pueda ser adaptado a situaciones educativas, clínicas, de orientación, rehabilitación y de otra índole relacionada.

El concepto de cociente de inteligencia

Cuando evaluamos la efectividad de la habilidad intelectual en un adulto, nos interesamos primeramente en dos tipos de comparaciones : qué lugar ocupa el individuo cuando se le compara con los de su edad, y qué lugar ocupa cuando se le compara con los que están en la cumbre de su desarrollo mental. Para facilitar tales comparaciones escogemos las técnicas que se emplean para convertir las puntuaciones brutas de la EIWA a puntuaciones a escala y de éstas a un cociente de inteligencia.

El Cociente de Inteligencia (CI) de la EIWA, ya sea basado en la Escala Verbal, en la de Manipulación o en la Total, se obtiene de la comparación directa de los resultados obtenidos en la prueba por el sujeto con aquellos de las personas de su grupo de edad cronológica. Es ésta, quizás, la fuente de información más significativa con relación a la habilidad mental del sujeto. La relación más significativa es la que lo compara con los compañeros de su misma edad. Se han preparado tablas de CI por separado de seis diferentes grupos de edad. Así que el CI de una persona de 22 años se basa en un grupo de sus coetáneos; el CI de un sujeto de 50 años se basa en las personas de ese mismo grupo de edad.

La ventaja de asignar CI de acuerdo con los grupos de edad adquiere mayor importancia si consideramos la variedad de experiencias que puede haber en la vida de las personas de diferentes edades. La cultura en la cual crecen las personas que tienen 18 años hoy día tiene muy poca semejanza con aquella que influyó en los que hoy tienen 60 años. La proporción de la población que ha mejorado su educación formal ha aumentado progresivamente en las últimas décadas. Por el contrario, las personas de mayor edad, por haber vivido más, tienen un cúmulo mayor de experiencias y han tenido más oportunidades de aprender y desarrollarse. Algunas influencias favorecen más a los adultos jóvenes y otras favorecen más a los de mayor edad.

Por las consideraciones anteriormente expuestas lo más apropiado sería comparar un sujeto con personas de edad similar a la suya. Un sujeto de 60 años que tenga un CI de 120 representa en su grupo de edad el mismo grado de superioridad mental que un sujeto de 18 años que tenga un CI igual (120) en su grupo. En este sentido --posición relativa dentro de su mismo grupo de edad-- el CI tiene un sólo significado, independientemente del grupo de edad en que se base.

Desde otro punto de vista, el significado de un CI es diferente en los distintos grupos de edad. En algunas edades se necesita más habilidad para obtener el mismo CI que en otras. Un número igual de respuestas correctas representan un CI más alto para un sujeto que pertenezca a un cierto grupo de edad que para otro que pertenezca a un grupo más joven. La evaluación de los resultados obtenidos en la prueba por un sujeto, independientemente de su edad, es importante para investigar el desarrollo de las habilidades mentales o cuando se necesite hacer aplicaciones prácticas en situaciones sociales, industriales y clínicas. Estas situaciones demandan el uso de una medida uniforme en la prueba que tendría el mismo significado para todos los sujetos. En este caso, el tipo de medida que se utiliza es la puntuación a escala.

Las puntuaciones a escala de la EIWA se basan en los resultados que obtuvo en la prueba un grupo de referencia seleccionado al efecto. La evidencia tiende a indicar que los adultos están en el máximo de su desarrollo intelectual entre los 20 y 34 años de edad. Se decidió usar como grupo de referencia a los sujetos de la muestra de estandarización incluídos en esas edades. En otros países en los que se ha llevado a cabo investigaciones similares a ésta también se ha usado ese mismo grupo de edad como grupo de referencia. La distribución de las puntuaciones de los 376 hombres y mujeres que comprendían este grupo fue la base para derivar las puntuaciones a escala. La tabla de equivalencias de puntuaciones brutas a escala, derivada de esta distribución, se aplica a todas las edades por igual. Una determinada puntuación bruta equivale siempre a la misma puntuación a escala no importa la edad del sujeto. Esta relación nos permite comparar los resultados obtenidos por una persona de cualquier edad con las del grupo de referencia. Las puntuaciones a escala son de gran utilidad para la investigación de la estabilidad o declinación de las habilidades según las miden las diferentes subpruebas, así como las escalas Verbal, Manipulación y Total.

La desviación del C I

El concepto original de C I envuelve una proporción entre la edad mental y la edad cronológica. En el caso de los niños la edad mental tiene algún significado; sin embargo, muchos psicólogos creen que esta medida se presta a confusión. En el caso de los adultos, la edad mental no tiene utilidad. Por esto la EIWA se diseñó para producir una desviación del C I al igual que la Wechsler Adult Intelligence Scale (WAIS) y la Wechsler Intelligence Scale for Children (WISC).

El método que se usó para computar la desviación del C I fue el de establecer un C I de 100 como promedio de las escalas Verbal, Manipulación y Total para cada grupo de edad. La desviación estándar se fijó a 15 puntos de C I para cada grupo de edad y se obtuvo a base de la distribución de las sumas de las puntuaciones a escala. De manera que la distribución de los C I tiene un promedio de 100 y una desviación estándar de 15 para cualquier grupo de edad.

Una ventaja importante de este método es que mantiene constante la variabilidad del C I de un grupo de edad a otro. Un C I específico tiene siempre el mismo significado en cuanto a la posición relativa que ocupa un sujeto en su grupo de edad. Un C I de 115 cae en la percentila 84 en las edades de 16, 35 ó 60 años, un C I de 90 está exactamente a dos tercios (2/3) de una desviación estándar bajo el promedio de ese grupo de edad, no importa la edad en la cual se haya obtenido el 90. El 50% de cada grupo de edad tendrá un C I entre 90 y 110. Esta constancia en la variabilidad facilita la interpretación del C I.

Organización de la prueba

La EIWA, como la WAIS, consiste de once subpruebas. Seis de ellas están agrupadas en la Escala Verbal y las otras cinco constituyen la Escala de Manipulación; las once se combinan para hacer la Escala Total.

Escala Verbal

Información
Comprensión
Aritmética
Analogías
Repetición de dígitos
Vocabulario

Escala de Manipulación

Dígito-Símbolo
Dibujos para completar
Diseños con cubos
Ordenamiento de dibujos
Composición de objetos

En términos generales existe mayor correlación entre las subpruebas de la Escala Verbal que entre éstas y las de la Escala de Manipulación. Lo mismo ocurre con las subpruebas de Manipulación en relación con las Verbales.

La Escala de Inteligencia Wechsler para Adultos es una traducción y adaptación de la WAIS estandarizada en la población puertorriqueña. Aquellos que tienen experiencia en el uso de la WAIS encontrarán que la EIWA es familiar, aunque también encontrarán muchos cambios. El profesional que use la EIWA deberá estudiarla cuidadosamente, como si fuese una prueba completamente distinta. Se hicieron cambios en el contenido del examen y en las instrucciones para la administración y valoración de la prueba. Algunos items de las diferentes subpruebas de la WAIS fueron eliminados; en otros se adoptaron formas alternas o se añadieron items completamente nuevos.[1] A continuación se presenta un resumen que ilustra los cambios a que nos referimos.

---

[1] En la subprueba de vocabulario, por ejemplo, los criterios usados para seleccionar los items fueron la universalidad y la frecuencia de los mismos. El orden de dificultad, (nivel de dificultad progresivo) se estableció a base de los resultados de un estudio especial de los items de esta subprueba.

Subpruebas	Núm. de items en WAIS	Núm. de items retenidos	Núm. de items con formas alternas	Núm. de items nuevos	Total núm. de items
Información	29	13	4	15	32
Comprensión	14	8	3	4	15
Aritmética	14	4	9	3	16
Analogías	13	5	2	8	15
Repetición de dígitos					
Directa	7	7	0	0	7
Inversa	7	7	0	0	7
Vocabulario	40	19	0	21	40
Dígito Símbolo*	90				110
Dibujo para completar	21	6	4	15	25
Diseños con cubos	10	8	1	1	10
Ordenamiento de dibujos	8	3	4	3	10
Composición de objetos	4	4	0	0	4

---

* Se usaron 6 pares de Dígito-Símbolo en vez de 9.

CAPITULO 2

ESTANDARIZACION DE LA ESCALA

La muestra de estandarización

El procedimiento más seguro para obtener una muestra representativa de una población consiste en obtener al azar una muestra no restringida de esa población. Este fue básicamente el método utilizado para obtener la muestra de la EIWA.

El Departamento del Trabajo proveyó una muestra al azar de áreas por unidades de vivienda, (en segmentos de seis unidades en la zona urbana y doce en la zona rural) en la cual se controlaron los factores de zona urbana o rural y las cinco regiones en que se dividió la isla. Facilitó además una lista de los residentes de 14 años o más de esas unidades de vivienda. Luego se obtuvo una muestra al azar de los ocupantes de estas viviendas. Esa muestra era restringida en el sentido de que el número de familiares que aparecen en ella se limitó arbitrariamente. En algunas ocasiones se aceptó el incluir un padre y su hijo pero no se aceptó incluir parejas de hermanos que vivían en la misma casa. En muchos sectores de la isla existen concentraciones de parientes que viven en relativa proximidad. Cualquier método de muestreo que se use inevitablemente resultaría en la inclusión de algunas de estas personas que tienen relación sanguínea.

El procedimiento utilizado consistió en darle a cada examinador una lista de sujetos incluídos en la muestra. La lista tenía el nombre, dirección, edad, ocupación, sexo y zona de residencia de cada sujeto. El examinador localizaba al sujeto y le administraba el examen inmediatamente o concertaba una cita para administrárselo luego. Generalmente se trataba de llevar a cabo el examen en las escuelas públicas de la localidad.

En algunas ocasiones el sujeto no se podía localizar, o no se lograba persuadirlo para que tomara el examen. En este caso, se seleccionaba un sustituto que fuese similar en edad, instrucción, ocupación y zona de residencia. La selección se hacía utilizando la muestra secundaria designada para este propósito.

El número total de sustituciones ascendió al 17% de la muestra. Solamente el 4% se perdió debido a la negativa del sujeto a cooperar. Los restantes se perdieron por razones tales como muerte, enfermedad física o mental, cambio de residencia sin notificación, emigración y otras. A pesar de las sustituciones antes mencionadas la muestra final resultó prácticamente igual a la original. De hecho, es probable que la muestra sea más representativa de la población del 1965 que de la población del Censo del 1960. Hay dos posibles fuentes de propensión. Primero, cuando un sujeto de la muestra primaria no podía examinarse y no se conseguía en la muestra secundaria un sustituto que tuviera características exactamente iguales. En este caso se trataba de conseguir un sujeto que pudiese reducir cualquier discrepancia entre la muestra y el Censo. En segundo lugar, al final del período de exámenes 49 sujetos quedaron sin examinar: estos incluían 29 urbanos y 20 rurales. Esta discrepancia resultó en sentido opuesto al que se esperaba.

La muestra de estandarización representa la población de Puerto Rico muy adecuadamente. Para ilustrar sus propiedades se presentan algunas tablas en las cuales los números y las proporciones de la muestra se comparan con las proporciones del Censo de 1960. Se notará que ocurren ciertas diferencias especialmente en relación con la instrucción. Las diferencias señalan la dirrección general de la tendencia poblacional en Puerto Rico.

Descripción y naturaleza representativa de la muestra

Las Tablas 1-5 resumen el material de comparación acerca de la naturaleza representativa de la muestra de estandarización. Cada tabla muestra tres tipos de porcentajes. Primero aparece el por ciento de población de acuerdo con el Censo del 1960. En segundo término está el por ciento de las mismas categorías (edad, sexo, instrucción, etc.), según aparecen en la muestra original. El Departamento del Trabajo proveyó la información necesaria para calcular estos por cientos cuando se obtuvo la muestra. En tercer lugar se presentan los por cientos por categorías de la muestra de estandarización que fue realmente examinada. Estos se calcularon a base de la información personal que se obtiene en el examen.

Debido a que varias de las características de la población de Puerto Rico están cambiando rápidamente es de esperarse que la muestra obtenida difiera un poco del Censo de 1960. Estas diferencias, sin embargo, deben seguir las tendencias conocidas de esta población. Ello es cierto en la mayoría de los casos. Las diferencias que se encuentran están en la misma dirección de los cambios que se sabe ocurren actualmente en la población puertorriqueña.

La Tabla I presenta la distribución de la muestra por edad y sexo. La Tabla 2 revela que la muestra incluye una proporción adecuada de hombres y mujeres en cada una de las regiones. La Tabla 3 indica que la muestra es representativa en términos de zonas de residencia (urbana-rural). La Tabla 4 describe la naturaleza representativa de la muestra en términos de categorías ocupacionales y hace evidente que existe un grado de concordancia apropiado entre la población del Censo del 1960 y la muestra final. La Tabla 5 indica que la muestra es razonablemente representativa en términos de logros educativos.

Tanto en la estandarización de la WAIS como en la de la EIWA se consideraron las mismas variables. Solamente se introdujeron algunos cambios menores: no se incluyeron los sujetos que presentaban condiciones físicas severas, disturbios emocionales y lesiones cerebrales; solo se aceptaron personas cuyo vernáculo fuese el español. El factor de la raza no fue considerado. A continuación se presenta una discusión de las variables que se consideraron.

I. Edad. Se establecieron normas para cada uno de los seis grupos de edad comprendidos entre los 16 y los 64 años. Los grupos de edad seleccionados corresponden a aquellos usados en la WAIS, con excepción del primer grupo. En la EIWA los dos primeros grupos de la WAIS, 16-17 y 18-19, se combinaron en uno. La siguiente tabla muestra el número de sujetos en cada grupo.

Tabla I. Muestra de Estandarización por Edad y Sexo

Grupo de edad	% Censo 1960		% Muestra Obtenida 1965		% Examinado	
	Masculino	Femenino	Masculino	Femenino	Masculino	Femenino
15	2.34	2.39	3.23	2.80	0.80	0.71
16-19	7.87	7.94	8.08	9.44	9.49	9.94
20-24	6.43	7.46	5.52	7.82	5.32	8.16
25-34	9.69	11.54	8.76	10.62	8.61	11.27
35-44	9.39	10.24	8.92	10.96	9.23	11.18
45-54	7.81	7.04	6.97	6.54	6.57	7.28
55-64	5.10	4.77	5.61	4.68	6.39	5.06
Total	48.63	51.38	47.09	52.86	46.41	53.60
N			554	622	523	604

2. Sexo. La proporción de hombres y mujeres incluída en cada grupo de edad se asemeja en gran medida al Censo del 1960.

3. Región Geográfica. El Departamento del Trabajo indentificó cinco regiones geográficas. Esta misma clasificación se usó en la muestra de la EIWA. La Región I la constituyen los pueblos de la sección noroeste; la Región II, los del sureste; la Región III, los pueblos del centro hacía la costa norte; la Región IV, los pueblos del oeste y la V la comprenden los pueblos del centro hacía la costa sur. El número de sujetos por región resultó ser proporcional a la población de cada una de éstas.

4. Residencia Urbana - Rural. La zona urbana la constituyen pueblos y ciudades de 2,500 o mas habitantes. Las comunidades de menos de este número de habitantes se definieron como zona rural. El por ciento de sujetos pertenecientes tanto a la zona urbana como a la rural fue muy similar al por ciento de cada zona dentro de la población del Censo.

Tabla 2. Muestra de Estandarización: Región por Zona de Residencia

	URBANA			RURAL		
Región	% Censo 1960	% Muestra 1965	% Examinado	% Censo 1960	% Muestra 1965	% Examinado
I	27.4	26.11	26.44	12.1	14.88	14.64
II	4.2	4.34	4.79	9.2	10.37	11.27
III	3.1	2.98	1.77	10.3	7.99	6.65
IV	4.9	4.68	5.59	11.0	11.56	11.98
V	7.9	9.18	8.78	10.0	7.91	8.08
Σ	47.6	47.29	47.37	52.4	52.71	52.61

Tabla 3. Muestra de Estandarización : por Edad, Sexo y Zona

		URBANA			RURAL		
Grupo de Edad	Sexo	% en Censo 1960	% en Muestra obtenida 1965	% Examinado	% en Censo 1960	% en Muestra obtenida 1965	% Examinado
15	M	.90	1.36	.27	1.44	1.87	.53
	F	1.00	1.62	.35	1.39	1.19	.35
16-19	M	3.17	3.66	4.08	4.70	4.42	5.41
	F	3.54	3.66	4.61	4.40	5.78	5.32
20-24	M	3.00	2.30	2.22	3.42	3.23	3.11
	F	3.67	3.74	3.37	3.79	4.08	4.79
25-34	M	4.96	4.17	4.44	4.73	4.59	4.17
	F	5.95	5.02	5.59	5.59	5.61	5.68
35-44	M	4.53	4.51	4.78	4.86	4.42	4.45
	F	5.06	6.46	6.12	5.18	4.51	5.06
45-54	M	3.47	3.06	2.75	4.34	3.91	3.82
	F	3.55	3.40	3.99	3.49	3.15	3.28
55-64	M	2.23	2.30	2.40	2.87	3.32	3.99
	F	2.52	2.04	2.40	2.25	2.64	2.66
Total		47.55	47.30	47.37	52.45	52.72	52.62

5. Ocupación. Las siguientes categorías ocupacionales se establecieron usando como base las agrupaciones del Censo.

00	-	Profesionales, técnicos, trabajos análogos
01	-	Agricultores y administradores de fincas
20	-	Administradores, oficiales y propietarios no de fincas

Tabla 4 A. Muestra de Estandarización - Por Sexo, Edad y Ocupación - Hombres
Categorías Ocupacionales

Grupo de Edad	I	II	III	IV	V	VI	VII y VIII	IX	X	XI	XII	XIII	XIV	XV
15														
a. % en el Censo 1960	--	--	--	.03	.01	.02	.02	.12	.02	--	3.02	1.60	.03	.03
b. % en Muestra obtenida 1965	--	--	--	--	.18	--	--	1.08	.18	--	--	--	--	5.42
c. % Examinado	--	--	--	--	--	--	--	--	.19	--	.96	--	.38	--
16-19														
a. % en el Censo 1960	.04	.03	.04	.42	.44	.54	.30	2.41	.51	--	6.02	4.84	.54	.10
b. % en Muestra obtenida 1965	--	--	--	.36	.72	1.08	.36	2.17	1.62	--	--	--	--	10.83
c. % Examinado	--	--	--	.76	.57	1.15	.57	1.72	2.29	--	10.33	.57	2.68	--
20-24														
a. % en el Censo 1960	.38	.09	.19	1.14	1.28	1.57	.54	2.31	1.82	--	1.22	2.00	.96	.11
b. % en Muestra obtenida 1965	.36	.18	--	1.44	1.99	1.99	.72	1.44	1.62	--	--	--	--	1.99
c. % Examinado	.19	.19	.19	1.34	1.34	1.91	.57	1.15	1.53	--	1.91	.38	.76	--
25-34														
a. % en el Censo 1960	1.47	1.29	1.02	2.25	2.49	3.06	1.11	3.03	1.14	--	.64	2.00	1.04	.15
b. % en Muestra obtenida 1965	1.08	.18	.72	2.17	3.43	3.97	1.44	1.99	1.81	--	--	--	--	1.62
c. % Examinado	1.53	.19	.38	2.10	3.36	3.82	1.53	1.91	2.10	--	--	1.19	1.15	--
35-44														
a. % en el Censo 1960	1.08	.58	1.69	1.98	2.65	2.60	1.15	3.60	1.32	--	--	2.01	.76	.10
b. % en Muestra obtenida 1965	.90	.72	2.71	1.62	4.51	1.99	.54	1.99	2.17	--	--	--	--	1.81
c. % Examinado	1.53	.57	1.34	2.29	4.40	3.44	.57	2.87	1.91	--	--	.19	.76	--
45-54														
a. % en el Censo 1960	.51	.81	1.64	1.45	2.01	1.66	.91	3.61	1.18	--	--	1.95	.54	.07
b. % en Muestra obtenida 1965	.36	.90	1.62	1.99	1.26	1.62	1.44	2.71	1.44	--	--	--	--	1.44
c. % Examinado	.19	.96	1.34	.76	1.72	2.87	.96	3.25	1.34	--	--	.19	.57	--
55-64														
a. % en el Censo 1960	.30	.70	1.02	.77	.90	.65	.47	2.20	.58	--	--	2.80	.27	.04
b. % en Muestra obtenida 1965	.18	1.62	1.26	1.08	1.62	.54	1.08	1.62	.90	--	--	--	--	2.17
c. % Examinado	.38	1.15	1.53	1.34	1.15	.38	.76	2.29	2.29	--	--	.57	1.91	--
Σ														
a. % en el Censo 1960	3.78	2.51	5.61	8.03	9.78	10.10	4.51	17.27	5.56	--	10.89	17.21	4.13	1.59
b. % en Muestra obtenida 1965	2.89	3.61	6.32	8.66	13.72	11.19	5.60	13.00	9.75	--	--	--	--	25.27
c. % Examinado	3.82	3.06	4.78	8.60	12.81	13.58	4.97	13.19	11.66	--	13.19	2.10	8.22	--

Tabla 4B. Muestra de Estandarización – Por Sexo, Edad y Ocupación – Mujeres

Categorías Ocupacionales

Grupo de Edad	I	II	III	IV	V	VI	VII y VIII	IX	X	XI	XII	XIII	XIV	XV
15														
a. % en el Censo 1960	--	--	--	.01	--	.01	.10	--	--	.16	2.73	1.72	.02	.03
b. % en Muestra obtenida 1965	--	--	--	--	--	--	--	--	--					5.31
c. % Examinado	--	--	--	--	--	--	--	--	--	--	.99	.33	.17	--
16-19														
a. % en el Censo 1960	.05	--	.01	.33	.02	.52	.64	.05	.01	2.39	5.78	5.24	.23	.08
b. % en Muestra obtenida 1965	--	--	--	.48	--	.32	.80	.48	--					15.92
c. % Examinado	--	--	--	.66	--	.17	.99	--	--	2.81	10.43	1.32	1.99	--
20-24														
a. % en el Censo 1960	.67	--	.03	1.17	.06	1.43	.70	.04	.02	5.80	.98	3.07	.35	.08
b. % en Muestra obtenida 1965	.32	--	--	1.61	--	1.29	.80	--	--					10.77
c. % Examinado	.50	--	--	1.32	--	.99	.66	.17	--	7.78	1.99	.83	.99	--
25-34														
a. % en el Censo 1960	1.28	--	.15	1.54	.10	1.90	1.07	.05	.02	12.48	.39	2.97	.40	.11
b. % en Muestra obtenida 1965	.32	.16	.32	2.09	--	3.22	.64	.16	.16					13.02
c. % Examinado	1.32	.17	.17	1.99	--	2.98	1.32	--	--	12.75	.17	.17	--	--
35-44														
a. % en el Censo 1960	.66	.01	.26	.87	.08	1.35	1.32	.05	.02	11.98	--	2.96	.26	.10
b. % en Muestra obtenida 1965	.80	--	.32	2.57	--	1.45	2.09	--	.16					13.34
c. % Examinado	.99	--	--	2.48	.17	1.99	2.15	--	--	12.91	--	--	.17	--
45-54														
a. % en el Censo 1960	.37	.01	.20	.41	.03	.58	.91	.04	.01	8.09	--	3.00	.11	.07
b. % en Muestra obtenida 1965	.48	--	.48	.16	--	1.13	.80	--	--					9.32
c. % Examinado	.83	.17	.17	.17	--	.83	1.16	--	.17	9.77	--	.17	.17	--
55-64														
a. % en el Censo 1960	.16	.02	.13	.13	.01	.23	.40	.03	.01	4.46	--	3.59	.04	.04
b. % en Muestra obtenida 1965	.16	--	--	--	--	.32	.64	.16	--					7.40
c. % Examinado	.17	--	--	.17	.17	.50	.50	--	--	7.62	--	.17	.17	--
Σ														
a. % en el Censo 1960	3.18	.05	.78	4.45	.30	6.01	5.14	.27	.09	45.36	9.89	22.55	1.41	.51
b. % en Muestra obtenida 1965	2.09	.16	1.13	6.91	--	7.72	5.79	.80	.32					75.08
c. % Examinado	3.81	.33	.33	6.79	.33	7.45	6.79	.17	.17	53.64	13.58	2.98	3.64	--

30 - Oficinistas y trabajos análogos
40 - Vendedores
50 - Artesanos, capataces, trabajos análogos
60 - Operarios y trabajos análogos
70 - Trabajadores de servicio doméstico y servicio protectivo, servicios de diversión y recreo, y asistentes en servicios profesionales y personales
80 - Obreros agrícolas y mayordomos
90 - Obreros
92 - Estudiantes
93 - Amas de casa
94 - Otros
98 - Desempleados
99 - Ocupación no informada

6. Instrucción. Se establecieron los siguientes ocho niveles a base de escolaridad (número de años de escuela que han completado) : 0 - 2 años, 3 - 4 años, 5 - 7 años, 8 años, 9 - 11 años, 12 años, 13 - 15 años, 16 ó más años. Se hicieron distribuciones de instrucción por separado de cada grupo de edad, por sexo.

Tabla 5. Muestra de Estandarización : por Edad, Sexo y Educación

Educación		0 - 2			3 - 4			5 - 7		
Grupos de edad	Sexo	Censo 1960	% en Muestra obtenida 1965	% Examinado	Censo 1960	% en Muestra obtenida 1965	% Examinado	% Censo 1960	% Muestra obtenida 1965	% Examinado
15	M	0.7	0.09	.00	0.31	0.51	0.09	1.13	1.28	0.18
	F	0.15	0.09	.00	0.25	.00	.00	1.00	1.36	0.44
16-19	M	0.69	0.34	0.35	0.86	0.77	0.98	2.44	2.13	2.40
	F	0.66	0.60	0.44	0.88	0.77	0.62	2.08	1.70	1.69
20-24	M	0.76	0.60	0.35	0.81	0.17	0.35	1.35	1.28	0.98
	F	1.01	0.51	0.62	1.11	0.51	0.53	1.56	2.13	1.95
25-34	M	1.44	1.11	1.33	1.60	1.53	0.98	1.66	1.45	1.51
	F	2.15	1.36	1.60	2.52	1.79	1.42	2.26	2.72	2.57
35-44	M	1.91	0.85	1.15	2.21	1.62	1.95	1.67	1.53	1.42
	F	2.79	2.21	2.22	2.75	2.81	3.28	1.93	1.87	1.86
45-54	M	2.21	2.13	2.04	2.23	1.87	1.69	1.50	1.45	1.33
	F	2.66	1.62	1.77	1.84	2.04	2.13	1.18	1.19	1.24
55-64	M	2.31	1.53	1.77	1.22	1.62	2.57	0.77	1.53	1.24
	F	2.61	1.96	2.48	0.84	1.28	0.98	0.63	0.60	0.62
SUMA		21.52	15.00	16.12	19.43	17.29	17.57	21.16	22.22	19.43

Cont. Tabla 5. Muestra de Estandarización: por Edad, Sexo y Educación

Educación		8			9 - 11			12		
Grupos de edad	Sexo	% Censo 1960	% Muestra obtenida 1965	% Examinado	% Censo 1960	% Muestra obtenida 1965	% Examinado	% Censo 1960	% Muestra obtenida 1965	% Examinado
15	M	0.45	0.68	0.18	0.29	0.68	0.27	.00	.00	.00
	F	0.53	0.51	0.18	0.45	0.85	0.18	.00	.00	.00
16-19	M	0.96	0.68	0.98	2.42	3.32	3.46	0.36	0.68	1.06
	F	0.82	0.77	0.53	2.77	4.17	4.53	0.53	1.19	1.51
20-24	M	0.47	0.34	0.35	1.37	1.28	1.15	1.01	1.36	1.33
	F	0.41	0.26	0.71	1.26	1.36	1.24	1.23	2.13	2.04
25-34	M	0.63	0.51	0.27	1.36	1.70	1.69	1.58	1.36	1.69
	F	0.68	0.43	0.71	1.29	1.11	1.15	1.39	2.30	2.31
35-44	M	0.85	1.02	0.80	0.90	1.53	1.24	0.95	1.45	1.33
	F	0.84	0.85	1.15	0.64	0.77	0.44	0.62	1.45	1.06
45-54	M	0.60	0.34	0.53	0.46	0.26	0.35	0.42	0.68	0.53
	F	0.42	0.51	0.62	0.34	0.43	0.62	0.28	0.26	0.27
55-64	M	0.27	0.34	0.18	0.19	0.17	0.35	0.17	0.17	.00
	F	0.22	0.26	0.44	0.15	0.17	0.27	0.13	0.17	0.09
SUMA		8.15	7.50	7.63	13.89	17.80	16.94	8.67	13.20	13.22

Cont. Tabla 5. Muestra de Estandarización : por Edad, Sexo y Educación

Educación		13 - 15			16		
Grupos de edad	Sexo	% Censo 1960	% Muestra Obtenida 1965	% Examinado	% Censo 1960	% Muestra obtenida 1965	% Examinado
15	M	.00	.00	.00	.00	.00	.00
	F	.00	.00	.00	.00	.00	.00
16-19	M	0.11	0.17	0.35	.00	.00	.00
	F	0.18	0.26	0.53	.00	.00	.00
20-24	M	0.50	0.43	0.62	0.12	0.17	0.18
	F	0.65	0.77	0.71	0.13	0.09	0.35
25-34	M	0.81	0.51	0.44	0.62	0.60	0.71
	F	0.81	0.43	0.89	0.50	0.51	0.62
34-44	M	0.40	0.34	0.62	0.52	0.60	0.71
	F	0.37	0.43	0.62	0.33	0.43	0.53
45-54	M	0.15	0.09	.00	0.24	0.17	0.09
	F	0.18	0.26	0.35	0.16	0.26	0.27
55-64	M	0.08	0.17	0.18	0.14	0.09	0.09
	F	0.12	0.09	.00	0.06	0.17	0.18
SUMA		4.36	3.95	5.31	2.82	3.09	3.73

Confiabilidad

Los coeficientes de confiabilidad de las subpruebas y de los C I Verbal, Manipulación y Total se presentan en la Tabla 6, por grupos de edad (16-19, 25-34, y 45-54). Estos tres grupos de edad fueron seleccionados como representativos de las edades que están incluídas en la muestra de estandarización.

En cada subprueba, excluyendo Repetición de Dígitos y Dígito-Símbolo, el coeficiente de confiabilidad se determinó computando la correlación entre puntuaciones en items pares y nones y corrigiendo el coeficiente de la subprueba completa mediante la fórmula Spearman-Brown. En la subprueba Repetición de Dígitos el coeficiente de correlación entre las puntuaciones de Repetición Directa y Repetición Inversa se corrigió para la subprueba completa.

Dígito- Símbolo es una subprueba de rapidez y la técnica de la prueba repartida (split-half) es inapropiada para computar su confiabilidad. No fue posible administrar una forma alterna de la subprueba Dígito - Símbolo a los sujetos en la muestra de estandarización después de haber sido sometidos a la prueba en su totalidad. Los sujetos pudieron haber reaccionado negativamente a la idea de tomar exámenes adicionales. Por consiguiente, se hizo un estudio especial para estimar la confiabilidad de esta subprueba.

Se escogieron tres grupos diferentes para hacer el análisis de correlación. El primer grupo consistió de 104 estudiantes exclusivamente de cuarto año de escuela superior. La edad de los sujetos fluctuó entre los 16 y los 20 años y ambos sexos estaban representados. El segundo grupo consistió de 50 adultos entre los 25 y 34 años de edad que cursaban el décimo, undécimo o duodécimo grado en una escuela superior nocturna. El tercer grupo estaba compuesto por 69 sujetos de ambos sexos cuyas edades fluctuaban entre los 35 y los 64 años. No se aceptó ningún sujeto que no tuviera instrucción primaria. La subprueba se administró dos veces con un receso de 5 minutos entre ambas. Para establecer la confiabilidad de esta subprueba se obtuvo una correlación entre el primer y segundo ensayo de cada uno de los tres grupos previamente descritos.

La desviación estándar que se obtuvo del primer grupo fue de 11.84 en el primer ensayo y de 12.27 en el segundo. Se usó la desviación estándar de 12.27 para calcular la confiabilidad que se hubiese obtenido de un grupo ilimitado usando la fórmula $\frac{\sigma_s}{\sigma_S} = \frac{\sqrt{1 - r_S}}{\sqrt{1 - r_s}}$.* Esta corrección no se aplicó a los dos grupos de mayor edad aunque en el caso del grupo de sujetos de 25 a 34 años pudo haberse usado. Sin embargo, la aplicación de esta fórmula en este caso hubiese resultado en un coeficiente inflado. El coeficiente de .87 de este grupo parece ser, pues, el más razonable.

Los coeficientes de confiabilidad de los CI Verbal, Manipulación y Total se computaron mediante una fórmula derivada de la fórmula de correlación entre dos sumas de puntuaciones igualmente ponderadas.**

Los valores de los coeficientes usados en la fórmula para Dígito -Símbolo son los siguientes: para los dos grupos más jóvenes se obtuvo valores arbitrarios de .71 y .77 calculando la confiabilidad a base de la mitad de la prueba y, para el grupo de mayor edad, se obtuvo una correlación de .71 eliminando de las puntuaciones la variabilidad que se atribuye a la edad.

Los coeficientes de confiabilidad que aparecen en la Tabla 6 deben ser cuidadosamente considerados por los clínicos al interpretar las puntuaciones de cada subprueba o las diferencias entre estas puntuaciones. A menor confiabilidad de una subprueba, menor seguridad se puede tener en los juicios hechos concernientes a la habilidad de un individuo. Los juicios de esta naturaleza deben hacerse con mucha precaución. Mientras más baja es la confiabilidad de las subpruebas mayor posibilidad hay de que la diferencia entre ellas se deba al azar en vez de a diferencias verdaderas en las habilidades del sujeto. Como se observa en la Tabla 6 esta precaución es menos necesaria cuando se usan las escalas totales (Verbal, Manipulación y Total) que tienen una confiabilidad alta.

---

* $\frac{\sigma_s}{\sigma_S} = \frac{\sqrt{1 - r_S}}{\sqrt{1 - r_s}}$, donde σ_s y r_s son la desviación estándar y el coeficiente de confiabilidad, respectivamente, de la amplitud restringida y σ_S y r_S son los números correspondientes de la amplitud no restringida.

** Kelley, T. L. Fundamentals of Statistics, Cambridge: Harvard University Press, 1947.

Tabla 6. Coeficientes de Confiabilidad y Errores Estándar de Medición (EEM)* de las Subpruebas y Escalas Totales

Subpruebas	r_{11}	EEM	r_{11}	EEM	r_{11}	EEM
Información	.93	.80	.94	.75	.94	.71
Comprensión	.79	1.33	.80	1.29	.84	1.20
Aritmética	.86	1.03	.86	1.10	.88	1.01
Analogías	.89	1.05	.93	.87	.91	.92
Repetición de dígitos	.66	1.55	.66	1.71	.65	1.54
Vocabulario	.95	.74	.95	.66	.96	.64
C I Escala Verbal	.97	2.80	.97	2.64	.97	2.64
Dígito - Símbolo **	.82	1.13	.87	1.13	.86	1.18
Dibujos para completar	.87	1.01	.91	.94	.92	.95
Diseños con cubos	.86	1.08	.89	1.05	.88	1.01
Ordenamiento de dibujo	.88	1.13	.89	1.04	.86	1.09
Composición de objeto	.76	1.40	.75	1.46	.80	1.32
C I Escala de Manipulación	.95	3.26	.96	2.85	.96	3.00
C I Escala Total	.98	2.28	.98	2.01	.98	2.07

---

*El EEM está expresado en unidades de puntuación a escala en el caso de las subpruebas y en C I en el caso de los tres totales (Verbal, Manipulación y Escala Total).

**Refiérase al texto si se interesa saber el método usado para obtener estos coeficientes.

En esta tabla se presentan los errores estándar de medición (EEM) de cada subprueba como otra evidencia de la estabilidad de las puntuaciones de la EIWA. Esta medida indica el margen de error de las puntuaciones de la prueba. De este modo un EEM de .80 en la subprueba de Información en las edades de 16-19 indica que hay dos oportunidades en cada tres de que la puntuación obtenida esté dentro de .80 puntuación a escala de la verdadera puntuación. Se puede estar bien seguro que la puntuación obtenida está dentro de tres errores estándar de la puntuación verdadera. Debe advertirse que existen diferencias notables en el valor de los EEM entre las distintas subpruebas. Por ejemplo, en las edades de 16 - 19 existe una probabilidad de dos a tres de que la puntuación de un sujeto en Vocabulario esté dentro de .74 puntuación a escala de su verdadera puntuación, pero en Repetición de Dígitos la misma probabilidad se aplica a un margen de error de ± 1.55 puntuación a escala. Es obvio que mientras más pequeño sea el EEM, mayor confianza se puede tener en una puntuación. Las diferencias entre las puntuaciones de Diseños con cubos y de Vocabulario, por ejemplo, se deben, con menor probabilidad, al azar que las diferencias entre Composición de objetos y Repetición de dígitos. Estos hechos demandan una precaución especial cuando se hacen comparaciones a base de diferencias entre perfiles y también dentro del mismo perfil.

Al leer los datos de la Tabla 6 el lector no debe confundirse por la discrepancia que se observa entre el valor de los EEM de las subpruebas y los EEM de los CI Verbal, Manipulación y Total. En las subpruebas los EEM se expresan en unidades de puntuación a escala; en las escalas totales se expresan en unidades de CI que son de interés y conveniencia para las personas que usan este examen. De manera que un error estándar de 3 en el CI Verbal de las personas que están dentro del grupo de edad 16-19 indica que existe una probabilidad de dos a tres de que el CI que se ha obtenido esté dentro de tres puntos del CI verdadero.

Intercorrelación de las subpruebas

Las intercorrelaciones de las once subpruebas de la EIWA se muestran en las Tablas 7, 8 y 9 y corresponden a los tres grupos de edad usados en el cómputo de los coeficientes de confiabilidad. Se incluyen también las correlaciones entre cada subprueba y las Escalas Verbal, Manipulación y Total y la correlación entre la Escala Verbal y la de Manipulación. El coeficiente se corregía para efectos de contaminación siempre que se correlacionaba una subprueba con una Escala Total de la cual esta subprueba formaba parte. (Ej. Vocabulario y Escala Verbal).

Tabla 7. Intercorrelaciones de la Subpruebas
Edades 16-19 Núm. Sujetos : 236

Subpruebas	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación Verbal	Puntuación de Manipulación	Puntuación Total
Comprensión	0.73													
Aritmética	0.64	0.60												
Analogías	0.72	0.63	0.57											
Repetición de dígitos	0.60	0.50	0.60	0.51										
Vocabulario	0.84	0.72	0.56	0.72	0.56									
Dígito - Símbolo	0.63	0.56	0.56	0.51	0.56	0.59								
Dibujos para completar	0.69	0.57	0.48	0.53	0.46	0.64	0.59							
Diseños con cubos	0.66	0.55	0.52	0.56	0.49	0.58	0.61	0.66						
Ordenamiento de dibujos	0.66	0.58	0.47	0.54	0.49	0.65	0.59	0.74	0.69					
Composición de objetos	0.55	0.52	0.44	0.45	0.41	0.51	0.53	0.68	0.68	0.61				
Puntuación * Verbal	0.86	0.76	0.70	0.75	0.64	0.82	0.68	0.68	0.67	0.68	0.58			
Puntuación de * Manipulación	0.76	0.66	0.58	0.62	0.57	0.71	0.67	0.79	0.78	0.78	0.73	0.78		
Puntuación * Total	0.86	0.76	0.68	0.73	0.65	0.81	0.72	0.76	0.76	0.76	0.67	--	--	
Promedio	9.97	9.24	10.03	9.70	10.53	9.41	10.67	10.53	10.26	10.71	10.58	58.88	52.75	111.64
Desviación Estándar	3.04	2.91	2.79	3.19	2.66	3.19	2.63	2.75	2.86	2.99	2.84	14.86	11.87	25.24

Correlación de Subpruebas con Puntuaciones a Escalas Verbal, de
Manipulación y Total sin Corregir para Efectos de Contaminación

Subpruebas	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación Verbal	Puntuación de Manipulación
Puntuación * Verbal	0.91	0.84	0.79	0.84	0.74	0.89	--	—	—	--	--	--	—
Puntuación de* Manipulación	--	—	--	--	—	—	0.78	0.87	0.87	0.87	0.83	--	—
Puntuación * Total	0.89	0.80	0.74	0.78	0.71	0.85	0.77	0.81	0.80	0.81	0.73	0.96	0.93

---

* Puntuación Verbal – Suma de 6 Subpruebas; Puntuación de Manipulación – Suma de 5 Subpruebas; Puntuación Total – Suma de las 11 Subpruebas. Coeficientes con estas variables en la parte principal de la tabla han sido corregidos para efectos de contaminación.

Tabla 8. Intercorrelaciones de las Subpruebas
Edades 25 - 34 - Núm. de Sujetos 224

Subpruebas	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación Verbal	Puntuación de Manipulación	Puntuación Total
Comprensión	0.77													
Aritmética	0.65	0.60												
Analogías	0.72	0.69	0.61											
Repetición de dígitos	0.61	0.61	0.70	0.59										
Vocabulario	0.82	0.78	0.62	0.73	0.65									
Dígito - Símbolo	0.73	0.67	0.64	0.58	0.64	0.70								
Dibujos para completar	0.71	0.66	0.55	0.62	0.60	0.66	0.69							
Diseños con cubos	0.63	0.60	0.64	0.56	0.64	0.64	0.70	0.69						
Ordenamiento de dibujo	0.72	0.67	0.57	0.63	0.55	0.66	0.76	0.76	0.68					
Composición de objetos	0.58	0.56	0.53	0.53	0.52	0.58	0.56	0.71	0.70	0.67				
Puntuación * Verbal	0.84	0.81	0.73	0.77	0.72	0.85	0.77	0.74	0.72	0.74	0.64			
Puntuación de* Manipulación	0.78	0.73	0.68	0.68	0.68	0.75	0.77	0.82	0.79	0.82	0.75	0.83		
Puntuación * Total	0.85	0.80	0.73	0.75	0.73	0.83	0.81	0.81	0.78	0.81	0.72	--	--	
Promedio	10.15	10.12	10.27	10.04	10.08	10.09	9.72	9.80	9.72	9.70	9.79	60.74	48.72	109.46
Desviación Estándar	3.03	2.86	2.95	3.20	2.93	3.06	3.12	3.11	3.15	3.10	2.99	15.40	13.44	27.62

Correlación de Subpruebas con Puntuaciones de Escalas: Verbal,
de Manipulación y Total sin Corregir para Efectos de Contaminación

Subpruebas	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación Verbal	Puntuación de Manipulación
Puntuación * Verbal	0.89	0.87	0.81	0.85	0.81	0.90	--	--	--	--	--	--	--
Puntuación* de Manipulación	--	--	--	--	--	--	0.86	0.89	0.87	0.89	0.84	--	--
Puntuación * Total	0.88	0.84	0.78	0.80	0.78	0.86	0.85	0.84	0.83	0.85	0.77	0.96	0.95

---

* Puntuación Verbal - Suma de 6 Subpruebas; Puntuación de Manipulación - Suma de 5 Subpruebas; Puntuación Total - Suma de las 11 Subpruebas. Coeficientes con estas variables en la parte superior de la tabla han sido corregidos para efectos de contaminación.

Tabla 9 Intercorrelaciones de la Subpruebas
Edades 45 -54 - Núm. de Sujetos 156

Subpruebas	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación Verbal	Puntuación de Manipulación	Puntuación Total
Comprensión	0.83													
Aritmética	0.63	0.64												
Analogías	0.70	0.68	0.57											
Repetición de dígitos	0.61	0.55	0.55	0.48										
Vocabulario	0.83	0.80	0.58	0.64	0.59									
Dígito - Símbolo	0.73	0.69	0.58	0.62	0.57	0.72								
Dibujos para completar	0.69	0.65	0.52	0.51	0.54	0.68	0.73							
Diseños con cubos	0.65	0.64	0.65	0.52	0.56	0.62	0.63	0.67						
Ordenamiento de dibujos	0.71	0.67	0.50	0.52	0.54	0.69	0.73	0.75	0.65					
Composición de objetos	0.56	0.58	0.50	0.46	0.57	0.58	0.63	0.72	0.65	0.64				
Puntuación * Verbal	0.87	0.85	0.69	0.72	0.64	0.82	0.78	0.71	0.72	0.72	0.64			
Puntuación de Manipulación*	0.78	0.75	0.63	0.61	0.64	0.77	0.78	0.83	0.74	0.80	0.76	0.83		
Puntuación * Total	0.86	0.83	0.69	0.69	0.67	0.83	0.82	0.80	0.77	0.79	0.72			
Promedio	10.07	10.07	10.55	10.15	9.92	10.14	7.38	7.78	8.99	7.94	8.64	60.90	40.72	101.62
Desviación Estándar	2.86	2.98	2.92	2.99	2.61	3.06	3.14	3.45	2.96	2.95	2.93	14.64	13.34	26.76

Correlación de Subpruebas con Puntuaciones de Escalas Verbal,
de Manipulación y Total sin Corregir para Efectos de Contaminación

	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación Verbal	Puntuación de Manipulación
Puntuación * Verbal	0.91	.90	0.79	0.81	0.74	0.89	--	--	--	--	--	--	--
Puntuación de* Manipulación	--	--	--	--	--	--	0.86	0.90	0.83	0.87	0.84	--	--
Puntuación * Total	0.89	0.86	0.75	0.75	0.72	0.87	0.86	0.84	0.81	0.83	0.77	0.96	0.95

---

* Puntuación Verbal - Suma de 6 Subpruebas; Puntuación de Manipulación - Suma de 5 Subpruebas; Puntuación Total - Suma de 11 Subpruebas. Coeficientes con estas variables en la parte superior de la tabla han sido corregidos para efectos de contaminación.

Esto se hizo computando la correlación entre la puntuación de una subprueba y la suma de las puntuaciones de las otras subpruebas de la escala correspondiente. Los coeficientes corregidos para efectos de contaminación se indican en la parte superior de las tablas y los que no fueron corregidos para fines de contaminación aparecen en la parte inferior de las tablas.

Derivación de las puntuaciones a escalas

Las puntuaciones a escala para cada subprueba se basan en un grupo de referencia de 376 sujetos de la muestra de estandarización entre las edades de 20 a 34 años. Los sujetos comprendidos en estas edades obtuvieron, en términos generales, las puntuaciones más altas entre los varios grupos de edad de la muestra. En cada subprueba la distribución de puntuaciones brutas del grupo de referencia se cambió a una escala con un promedio de 10 y una desviación estándar de 3. Esto se llevó a cabo por medio de una distribución de frecuencia cumulativa de las puntuaciones brutas y fijando cada puntuación percentil en el valor apropiado de la puntuación estándar en una curva normal teórica. Para cada sujeto de la muestra de estandarización las puntuaciones brutas de las subpruebas se transformaron en puntuaciones a escala usando como base el grupo de referencia. Por lo tanto las puntuaciones a escala hacen posible que se puedan establecer comparaciones entre los resultados individuales de los sujetos de cualquier grupo de edad con el grupo de referencia. Debe advertirse que este tipo de comparación se adapta solamente a las puntuaciones de las subpruebas. El método de obtener los C I se presta primordialmente para comparar el C I de un sujeto con el de sus contemporáneos.

Derivación de las Tablas de los Cocientes de Inteligencia

Las puntuaciones a escalas de las subpruebas se obtienen de una tabla que se basa en el grupo de referencia; sin embargo, los C I se han derivado para cada uno de los seis grupos de edad por separado. Se fijó de antemano un promedio de 100 y una desviación estándar de 15 para cada grupo de edad. Se computaron los promedios y las desviaciones estándar del total de las puntuaciones a escala de las seis subpruebas Verbales, del total de las cinco subpruebas de Manipulación y del total de las puntuaciones a escala de las once subpruebas. Esta información se presenta en la Tabla.10.

Como podrá observarse las seis desviaciones estándar de las sumas de las puntuaciones Verbales, y de las de Manipulación, así como las sumas de las puntuaciones de todas las subpruebas son muy similares. Con el fin de determinar si se sostenía la hipótesis de que no existen diferencias se decidió aplicar la prueba de Bartlett, para homogeneidad de la varianza a cada uno de los tres grupos de seis varianzas. En cada caso se sostuvo la hipótesis de que las diferencias entre las varianzas de los seis grupos de edad se debían a la casualidad. Las probabilidades que se obtuvieron de los valores de r^2 a base de la prueba de Bartlett fueron .59, .11 y .25 para las sumas de las puntuaciones Verbales, Manipulación y Total respectivamente. Por lo general, un valor de probabilidad mayor de .05 se considera adecuado para sostener la hipótesis de que no hay diferencias significativas entre las dos variables.

Tabla 10. Promedios Aritméticos y Desviaciones Estándar de las Sumas de las Puntuaciones a Escala: Verbal, Manipulación y Total, por Grupos de Edad

Grupo de Edad	N	Promedio aritmético	Desviación estándar	Promedio aritmético	Desviación estándar	Promedio aritmético	Desviación estándar
16-19	236	58.88	14.83	52.75	11.85	111.64	25.18
20-24	152	58.89	16.32	50.84	13.17	109.72	28.40
25-34	224	60.74	15.37	48.72	13.41	109.46	27.56
35-44	230	62.29	16.04	46.10	14.29	108.38	29.16
45-54	156	60.90	14.60	40.72	13.30	101.62	26.68
55-64	129	57.64	16.09	36.95	13.75	94.58	28.90

Se computó la desviación estándar de cada una de las tres Escalas totales (Verbal, Manipulación y Total) y se usaron en todos los grupos de edad. Las desviaciones estándar obtenidas fueron 15.51 en la Escala Verbal, 13.58 en Manipulación y 27.57 en la Escala Total.

Las Tablas de C I se prepararon fijando el promedio de las escalas totales en 100 y la desviación estándar en 15 en cada grupo de edad y transformando todas las puntuaciones en C I. Estas tablas aparecen en las páginas 104 a 115 de este Manual.

La Tabla 11 facilita la interpretación de los C I en términos de percentilas.

Tabla 11. Relación de los C I de la EIWA con Rangos Percentiles

Rango percentil	CI equivalente	Rango percentil	C I equivalente
99	135	50	100
97	128	40	96
95	125	30	92
90	119	25	90
80	113	20	87
75	110	10	81
70	108	5	75
60	104	3	72
		1	65

Interpretación de los cocientes de inteligencia

Los diagnósticos equivalentes a los distintos niveles de inteligencia se presentan en la Tabla 12. Esta clasificación se obtuvo estadísticamente basándose en la teoría de que ésta es la forma más precisa de definir niveles de inteligencia. No se pretende que estas delimitaciones necesariamente se ajusten a definiciones universalmente aceptadas, sin embargo, es conveniente que de fijarse otros límites se indique siempre su justificación estadística.

Tabla 12. Clasificaciones de Inteligencia

C I	Clasificación	% Incluído
130 ó más	Muy superior	2.2
120 - 129	Superior	6.7
110 - 119	Normal brillante	16.1
90 - 109	Promedio	50.0
80 - 89	Normal torpe	16.1
70 - 79	Fronterizo	6.7
69 ó menos	Deficiente mental	2.2

Correlación de la EIWA con la Prueba Otis

Se hizo un estudio especial para determinar la relación entre la EIWA y la prueba Otis, Superior Modelo A[1].

La muestra primaria que se obtuvo para la estandarización de la EIWA se dividió en dos partes a fin de facilitar el proceso de administración. de la prueba. La primera parte se realizó durante los primeros siete meses y la segunda durante los siguientes ocho meses. La muestra utilizada para el estudio especial de correlación se obtuvo de esta segunda parte de la muestra primaria. Los sujetos escogidos para este estudio fueron todos aquellos que fluctuaban entre las edades de 16 a 45 años y que tenían de 7 a 18 años de instrucción escolar. La muestra de la prueba Otis la constituyó un total de 205 sujetos que reunían esas características. La distribución de edad y nivel de instrucción de los sujetos se ilustra en la siguiente tabla.

Tabla 13. Edad y Nivel de Instrucción de la Muestra Otis

Grupo de edad	Frecuencia	%	Educación	Frecuencia	%
16 - 19	69	33.66	7-8	34	16.58
20- 24	41	20.00	9-10	54	26.34
25-29	27	13.17	11-12	82	40.00
30-34	24	11.71	13-14	15	7.32
35-39	22	10.73	15-16	10	4.88
40-45	22	10.73	17-18	10	4.88

La prueba Otis fue administrada por el mismo examinador y en el mismo sitio y fecha que la EIWA. Primeramente se administraba la EIWA y después de diez minutos, aproximadamente, se procedía a administrar la Otis. La administración de la Otis se llevó a cabo siguiendo los procedimientos estipulados en el Manual de Instrucciones.

---

[1] Esta es una traducción y adaptación hecha en Guatemala de la prueba en inglés Otis Quick Scoring Mental Ability Tests, de Arthur S. Otis. Fue estandarizada con población de escuela superior.

Para determinar la relación entre la EIWA y la Otis se realizó un estudio de correlación usando las tres puntuaciones totales brutas de la EIWA y la puntuación total bruta de la Otis.

Para establecer las comparaciones entre los promedios y la desviación estándar de ambos exámenes se obtuvo el promedio de las tres puntuaciones totales brutas de la EIWA y sus respectivas desviaciones estándar, así como también el promedio y la desviación estándar de la puntuación total bruta de la Otis. Los resultados se presentan en la Tabla 14.

Tabla 14. Coeficientes de Correlación entre la Prueba Otis y la EIWA

Prueba	r con Otis	Promedio	Desviación Estándar
Otis	--	24	12
EIWA			
Verbal	.86	96	32
Manipulación	.73	137	37
Escala Total	.85	233	64

N = 205

Se encontraron coeficientes de correlación de .86, .73 y .85 entre la puntuación total bruta de la Otis y los tres totales de la EIWA (Verbal, Manipulación y Total) respectivamente. Como se esperaba, la correlación entre la EIWA y la Otis resultó alta y positiva. Resultados similares halló Wechsler en sus estudios de correlación entre el WAIS y el Stanford-Binet.[1]

---

[1] Wechsler, David - Wechsler Adult Intelligence Scale Manual, The Psychological Corporation, New York, 1955, Page 21.

CAPITULO 3

CONSIDERACIONES GENERALES PARA LA ADMINISTRACION DE LA ESCALA

Instrucciones y procedimientos

En la administración de la Escala de Inteligencia Wechsler para Adultos (EIWA) deben seguirse fielmente las instrucciones de este Manual. Los procedimientos aquí prescritos se elaboraron a fin de producir una evaluación fiel de la habilidad mental. Estos procedimientos se ensayaron antes cuidadosamente durante la elaboracion de la Escala y se siguieron de manera estricta en la estandarización de la prueba. Por lo tanto, el obtener resultados válidos mediante el uso de la EIWA depende, por lo menos en parte, del seguir fielmente las instrucciones establecidas. Esta advertencia se aplica igualmente a los procedimientos de valoración. La alteración de los métodos indicados para acreditar puntos o calcular puntuaciones totales, puntuaciones a escala y cocientes intelectuales sólo producirían evaluaciones erróneas de la inteligencia del sujeto.

Los examinadores que están acostumbrados a emplear traducciones de las Escalas de Inteligencia Wechsler deben tener particular cuidado de no emplear los procedimientos de esas escalas al administrar la EIWA. El revertir al empleo de esos procedimientos es frecuente cuando dichos procedimientos se asemejan a los de la nueva Escala. Sin embargo, difieren en aspectos importantes y por lo tanto, debemos gardarnos cuidadosamente de no incurrir en esa falta.

Como en cualquier otra prueba psicológica, la administración correcta de la EIWA requiere un examinador competente, materiales organizados, un salón de pruebas apropiado y tiempo suficiente. El examinador deberá estar adiestrado especialmente, tanto en la administración de pruebas individuales en general, como en la administración de la EIWA en particular. Debe de practicar los procedimientos de la Escala antes de administrarla al sujeto. Deberá familiarizarse completamente con las instrucciones y condiciones especiales correspondientes a cada subprueba de la EIWA. Esto es necesario para que el examinador pueda seguir el Manual fácilmente y sin titubeos, anotar las respuestas con exactitud y responder apropiadamente a las reacciones del sujeto y a sus respuestas, cualesquiera que éstas fueren.

El examinador debe arreglar los materiales de la prueba convenientemente, de manera que pueda presentarlos sin confusión ni retraso cuando los necesite. Estos materiales deben mantenerse fuera de la vista del sujeto hasta que se necesiten en el transcurso de la prueba. Por otra parte, tratar de ocultar en forma ostentosa los materiales pueden provocar en el sujeto una reacción desfavorable.

El salón donde se administre la prueba debe estar bien iluminado, ventilado y libre de estímulos e interrupciones que distraigan la atención del sujeto. El mobiliario ha de arreglarse en forma tal que el examinador y el sujeto estén cómodos, que el sujeto pueda manipular libremente los materiales con que ha de bregar y que el examinador pueda presentar los materiales, observar la ejecución del sujeto y anotar convenientemente las respuestas.

Se debe disponer de tiempo suficiente para administrar la prueba, de manera que se establezca y mantenga buen "rapport" y que la administración de la prueba se desarrolle en forma cómoda y pausada.

El examinador debe esforzarse por obtener la cooperación del sujeto y mantener su motivación. Es aconsejable y necesario en muchos casos, hacer comentarios alentadores durante el examen. Los siguientes comentarios se consideran apropiados, dependiendo del contexto del ítem o de la subprueba: "Bueno, eso no le tomó mucho tiempo"; "Aquí tenemos algo diferente"; "Yo creo que usted encontrará esto interesante". Sin embargo, no es apropiado hacer un comentario como: "Eso está correcto"; "No, está malo"; "Muy bien"; inmediatamente después de una respuesta, ya que el sujeto podría esperar en ocasiones siguientes que se le informe sobre cuán correctas son sus otras respuestas y puede interpretar el no hacer comentarios como indicio de que ha fallado. De la misma manera, si se tiene que hacer un segundo ensayo de un ítem de demostración, no es prudente decir: "Ahora pruebe esto otra vez". El examinador debe rehusar dar información al sujeto acerca de si son o no acertadas sus respuestas, pero deberá asegurarse de mantener el esfuerzo y el interés del sujeto durante la prueba.

El orden para la administración de las subpruebas de la EIWA es el siguiente:

Subpruebas Verbales

1. Información
2. Comprensión
3. Aritmética
4. Analogías
5. Repetición de dígitos
6. Vocabulario

Subpruebas de Manipulación

7. Dígito - Símbolo
8. Dibujos para completar
9. Diseños con cubos
10. Ordenamiento de dibujos
11. Composición de objetos

Es muy deseable que se siga el orden indicado en la administración de las subpruebas para asegurar la aplicabilidad estricta de las normas a los resultados, ya que éste fue el orden que se utilizó con todos los sujetos de la muestra de estandarización. Habrá ocasiones en que será preferible variar el orden de presentación; por ejemplo, si un sujeto tiene un impedimento del habla el administrarle primeramente la subprueba de Diseños con cubos o la subprueba de Composición de objetos lo pondría en mejor disposición que si se empieza la prueba según el orden establecido. En términos generales, las modificaciones en el orden de la administración sólo deben hacerlas examinadores que tengan gran experiencia.

Además del Manual y la Hoja de anotaciones en la administración de la EIWA se usa un equipo que incluye lo siguiente:

1. Una tarjeta que contiene la lista de palabras de la subprueba de Vocabulario.

2. Una libreta de lomo cosido con espiral de alambre que contiene los ítems de la subprueba de Dibujos para completar y, los diseños para la subprueba de Diseños con cubos. El contenido de esta libreta está organizado de tal manera que facilita una presentación adecuada del material.

3. Una libreta de lomo cosido con espiral de alambre que contiene las tarjetas para la subprueba de Ordenamiento de dibujos. Cada grupo de tarjetas aparece en un bolsillo por separado.

4. Nueve bloques rojos y blancos de forma cúbica para la subprueba de Diseños con cubos; éstos también se usan para administrar el ítem número 1 de la subprueba de Aritmética.

5. Cuatro cajas que contienen las partes de los ítems de la subprueba de Composición de objetos.

6. Una pantalla de cartón en forma de biombo en la que se muestra cómo deben arreglarse las partes para presentarlas al sujeto en la subprueba de Composición de objetos. El otro propósito de esta pantalla es esconder las partes de la subprueba de la vista del sujeto mientras se arreglan para presentárselas.

7. Una clave para la valoración de la subprueba de Dígito-Símbolo.

La Hoja de respuestas facilita la anotación de las respuestas y cualquiera otra información pertinente acerca del sujeto y su conducta durante el examen. En algunas de las subpruebas, v.g. la de Información y la de Dibujos para completar, los ítems pueden valorarse mientras el sujeto da las respuestas. Sin embargo, en las subpruebas de Comprensión, Analogías, Vocabulario y Ordenamiento de dibujos, es imperativo que el examinador anote exactamente las respuestas tal como las da el sujeto. El anotar las respuestas del sujeto al pie de la letra permite una consideración más completa de ellas y una evaluación posterior.

Al administrar las subpruebas, el examinador siempre debe leer las instrucciones y las preguntas directamente del Manual. De no hacerlo así, podría cambiar la fraseología y apartarse de los procedimientos establecidos. Cuando sea necesario, el examinador puede repetir las instrucciones o las preguntas al sujeto (excepto en la subprueba de Repetición de dígitos), pero no debe parafrasearlas ni adornarlas, excepto cuando así se indique en las instrucciones de cada subprueba.

Se deben leer las instrucciones y las preguntas clara e inteligiblemente. Las fallas que el sujeto pueda tener en entenderlas nunca deben ser causadas por la enunciación defectuosa del examinador.

En las subpruebas Verbales de Información, Comprensión, Analogías y Vocabulario, si una respuesta está incompleta o no está clara, el examinador puede pedirle al sujeto que la explique más. Esta técnica, a la cual nos referimos frecuentemente como repregunta, puede mejorar tanto la exactitud de la valoración como la comprensión cualitativa de las respuestas. Sin embargo, le impone al examinador la responsabilidad de estar bien familiarizado con los criterios de valoración y de usar hábilmente la repregunta.

Se debe evitar el excesivo sondeo del sujeto, presionarlo para que responda o ayudarlo haciéndole preguntas que le sugieran las respuestas adecuadas. De igual manera, no se deben usar procedimientos tales como deletrear las palabras de las preguntas ni alterar la fraseología de un ítem. Es necesario que la repregunta sea neutral y que no le sugiera al sujeto que sus respuestas se pueden mejorar. Tampoco se debe convertir ésto en un procedimiento para presionar al sujeto con el propósito de que llegue al límite de sus conocimientos. Sin embargo, es apropiado usar expresiones tales como: "¿Quiere explicarme un poco más?"; "Por favor, explíquese más", "¿Podría usted repetir eso?"; o "Continúe". Esto le permite al sujeto aclarar sus aseveraciones, de modo que puedan ser propiamente evaluadas. Cuando se hagan repreguntas, el examinador debe anotar el símbolo (?) e inmediatamente anotar la respuesta que el sujeto da a la repregunta. Así, en la evaluación subsiguiente de las respuestas, estará claro cuáles de ellas fueron espontáneas y cuáles fueron provocadas por las repreguntas.

Si el sujeto corrige o aclara una respuesta, ya sea espontáneamente o como resultado de las repreguntas, se le debe dar crédito por la respuesta corregida, de acuerdo con los criterios de valoración. Sin embargo, el ítem debe valorarse en cero, si la elaboración revela que el sujeto tiene un concepto completamente erróneo de la solución.

En aquellas subpruebas para las cuales no se fijan límites de tiempo, el examinador debe darle tiempo adecuado al sujeto para que formule y exprese sus respuestas. Lo que se considere tiempo adecuado variará de acuerdo con el nivel de actividad del sujeto y de la atención que éste demuestre , y tendrá que ser estimado por el examinador en cada caso. En la mayoría de los casos, de 10 a 15 segundos es suficiente tiempo para que el sujeto comience a dar su respuesta. Si el sujeto no respondiera, aún después de haberse hecho la invitación apropiada, o después de haberle repetido la pregunta, el examinador le anotará una puntuación de cero a ese ítem y proseguirá con el siguiente.

En aquellas subpruebas donde se prescribe un límite estricto de tiempo y donde en algunos casos se da crédito adicional por rapidez, el examinador debe empezar a medir el tiempo al finalizar la presentación de cada ítem. En estos casos se tiene que usar un cronómetro. Si es necesario repetir un ítem, aclarar una respuesta o hacer alguna advertencia después de la presentación inicial del ítem, dichas instrucciones adicionales se cuentan como parte del tiempo asignado para contestar el ítem.

Los criterios de suspensión de cada subprueba aparecen junto con las instrucciones específicas para la administración de la misma. La mayor parte de las veces cierto número de fallas consecutivas (respuestas valoradas en 0) constituyen una base para suspender esa prueba y empezar la próxima. En algunos casos, todos los ítems se administran a todos los sujetos, y en una subprueba en particular, Dígito -Símbolo, el límite de tiempo determina el número de ítems que habrán de administrarse. En las subpruebas de la primera clase siempre que el examinador tenga duda de si se ha cumplido o no con el criterio de suspensión, debe continuar presentándole los ítems al sujeto hasta que esté seguro de que una evaluación posterior de las respuestas verificará la existencia del número de fallas consecutivas requeridas. En otras palabras , es un procedimiento más seguro examinar al sujeto en exceso que hacer lo contrario.[1]

Las instrucciones para valorar las subpruebas aparecen también en las secciones de administración en este Manual. En todas las subpruebas de Manipulación y tres de las subpruebas Verbales, la valoración de las respuestas es completamente objetiva. Sin embargo, la valoración de las respuestas de las subpruebas de Comprensión, Analogías, y Vocabulario, requiere una considerable participación del juicio del examinador. Es imposible anticipar todas las posibles respuestas que podrán darse a los ítems de esas subpruebas o hacer una lista de ellas. Se proveen, en una sección separada del Manual, criterios generales de valoración bastante amplios y ejemplos de respuestas valoradas en 2, 1 y 0. El examinador debe familiarizarse enteramente con estos criterios y aprender los significados de las palabras y de los conceptos incluídos.

La labor oficinesca del examinador

La administración y valoración de la EIWA envuelve cierta labor de índole oficinesca. El examinador deberá reconocer la importanica de realizar estas tareas con cuidado y precisión. No se debe presumir que el ser un profesional garantiza la eficiencia en este trabajo, que es mayormente de tipo oficinesco. Al llenar la Hoja de anotaciones el examinador debe tomar en consideración lo siguiente:

1. Al valorar cada ítem haga sus anotaciones precisas y legibles.

2. En los ítems que conllevan bonos adicionales por tiempo, anote el tiempo correctamente para el ítem en cuestión y acredite el bono correctamente.

3. En las subpruebas de Información, Comprensión, Aritmética y Vocabulario en las que puede que no tenga que administrar los primeros ítems, acredite los puntos para esos items.

4. Al computar las puntuaciones brutas totales de las subpruebas coteje la suma de los puntos obtenidos en los ítems.

---

[1] Se hizo un análisis de dificultad de ítems basado en la muestra de estandarización, (1,127 sujetos). Este análisis reveló que los ítems siguen un orden ascendente de dificultad, esto es, aparecen de menor a mayor dificultad. Existen algunas excepciones a esta regla, pero son de menor importancia.

5. Asegúrese de que la puntuación bruta de cada subprueba se transfiera al espacio correspondiente en la Sección de resumen de los resultados que aparece en la portada de la Hoja de anotaciones.

6. Evite errores al convertir las puntuaciones brutas a las puntuaciones a escala y también al convertir éstas en C I. Repita los pasos envueltos en este proceso para corregir posibles errores en la lectura de las tablas.

7. Verifique la edad del sujeto comparando la edad informada por el sujeto con la diferencia entre la fecha de administración de la prueba y la fecha de nacimiento, o consultando récords confiables.

8. Revise toda la información transferida, las computaciones y la conversión de puntuaciones para más exactitud.

Conversión de las puntuaciones brutas a puntuaciones a escala

La puntuación bruta en cada subprueba es la suma del número de puntos obtenidos en cada ítem. Estas puntuaciones brutas se transfieren a la sección de resumen de los resultados que se provee en la portada de la Hoja de anotaciones. *Hacia la izquierda de ésta hay una Tabla* de equivalencias a puntuaciones a escala (que aparece también en la página 103 de este Manual) que se usa para convertir las puntuaciones brutas en puntuaciones a escala de todos los sujetos, independientemente de su edad y sexo. La puntuación bruta que un sujeto obtiene en una subprueba se localiza en la columna de la tabla correspondiente a esa subprueba. El examinador podrá localizar la puntuación a escala correspondiente a cierta puntuación bruta leyendo horizontalmente hacia la izquierda o hacía la derecha de la columna en la Tabla. Esta puntuación a escala se anota en el espacio correspondiente en la Sección de resumen de los resultados a la derecha de la puntuación bruta. Al terminar este procedimiento con todas las subpruebas, la Sección de resumen de los resultados muestra una columna de puntuaciones brutas y al lado de ésta, una de puntuaciones a escala. De aquí en adelante no es necesario usar las puntuaciones brutas, pues las puntuaciones a escala son más significativas y permiten tratamiento aritmético.[1]

---

[1] Se provee a la derecha de esta columna un espacio adicional que el examinador puede usar para colocar las desviaciones de puntuación del sujeto en cada subprueba o cualquier otra anotación que se quiera hacer.

La Escala Verbal es la suma de las puntuaciones a escala de las seis subpruebas y la de Manipulación se obtiene sumando las puntuaciones a escala de las cinco subpruebas de Manipulación. La suma de la Escala Verbal y la de Manipulación constituyen la Escala Total que está basada en once pruebas.

En algunas ocasiones no es posible usar algunas de las subpruebas bien sea porque se considere que el resultado obtenido no es válido, o por alguna limitación del cliente. En estos casos hay que proratear para derivar el C I si se omite una de las subpruebas Verbales, la suma de las puntuaciones a escala de las cinco restantes se multiplica por seis quintos (6/5) para obtener la puntuación a escala Verbal . En caso de que sea una subprueba de ejecución la que se omita, se multiplica la suma de las puntuaciones a escala de las otras cuatro subpruebas por cinco cuartos (5/4) para determinar la puntuación a escala de Manipulación. Es posible proratear cuatro subpruebas Verbales multiplicando la suma de sus puntuaciones a escala por 3/2 y también proratear tres subpruebas de Manipulación por el factor de 5/3, pero esto no es recomendable. Las puntuaciones de la Escala Total deben siempre basarse en por lo menos cinco subpruebas Verbales y cuatro de Manipulación. Debe observarse que sólo se proratea para obtener las puntuaciones Verbales y de Manipulación por separado, pero nunca se proratea la puntuación de la Escala Total . Después que se obtienen en esta forma las puntuaciones de las Escalas, Verbal y Manipulación éstas se suman para obtener la puntuación de la Escala Total, como se hace regularmente. Las puntuaciones que han sido prorateadas deben identificarse con el símbolo PRO al margen de la sección de Resumen de la Hoja de anotaciones.

Método para determinar los cocientes de inteligencia

La Tabla 16, que aparece en las páginas 104-115, se usa para transformar las puntuaciones a escala (Verbal , Manipulación y Total) en C I. Esta tabla consiste de seis grupos de tablas, una para cada grupo de edad. Los C I de las escalas Verbal , Manipulación y Total , se obtienen de páginas contiguas (inmediatas) de manera que el examinador pueda determinar fácilmente los tres C I de un sujeto al consultar el grupo de tablas correspondientes a su grupo de edad.

La edad que se considera es el último año cumplido por el sujeto. Por ejemplo, si la persona tiene 34 años con 10 meses cuando toma la prueba, su edad se considera 34 años y su C I se encontraría en la serie de tablas correspondiente al grupo de edad 25-34.

El procedimiento que se usa es el siguiente: busque en las tablas (Tabla 16) el grupo de edad a que pertenece el sujeto. En el lado izquierdo de la página de la izquierda se encontrarán puntuaciones de la Escala Verbal con sus correspondientes C I. En el lado derecho de esa misma página, se encontrarán las puntuaciones de la Escala de Manipulación con sus correspondientes C I. La página a mano derecha contiene las puntuaciones de la Escala Total y sus respectivos C I. Localice la puntuación Verbal del sujeto en la columna de puntuaciones verbales y anote el C I verbal correspondiente en el espacio que se provee en la sección de Resumen de los resultados de la Hoja de contestaciones. En esta forma anote también los C I que se obtengan de las Escalas de Manipulación y Total .

A continuación se presenta un ejemplo de cómo determinar los C I de un sujeto que tenía 42 años y 7 meses cuando tomó la prueba. Su puntuación en la Escala Verbal (la suma de las puntuaciones a escala de las seis subpruebas Verbales) fue 68; su puntuación en la Escala de Manipulación (la suma de las puntuaciones a escala de las cinco subpruebas de Manipulación) fue 116. Usando la Tabla 16 y tomando en consideración la edad del sujeto (35-44), se observa que éste obtuvo un C I de 108 en la Escala Verbal, de 103 en la de Manipulación y de 106 en la Escala Total de la prueba.

INSTRUCCIONES PARA LA ADMINISTRACION DE LAS SUBPRUEBAS

INFORMACION

INSTRUCCIONES Comience siempre con la pregunta número 4. Si el sujeto falla en las preguntas 4 y 5 administre entonces las preguntas de la 1 a la 3 inclusive. Si el sujeto falla en las preguntas de la 1 a la 3 inclusive, administre las preguntas 6 y 7 y si falla en todas, suspenda la subprueba.

Si una de las respuestas del sujeto está incompleta o no está clara, el examinador puede pedirle que se explique más ampliamente. Sin embargo, el examinador no debe explicar el significado de las palabras, ni alterar la fraseología de las preguntas.

SUSPENDA la subprueba luego de 7 fallas consecutivas.

VALORACION Déle un punto por cada una de la respuestas correctas. No se dan puntos por contestaciones que son aceptables sólo en parte. Déle 3 puntos a los sujetos a quienes no haya que administrársele las preguntas de la 1 a la 3 inclusive.

A continuación se citan las partes esenciales de las respuestas aceptables. En los casos en que se anotan posibles respuestas (las cuales aparecen separadas por puntos suspensivos), basta con que el sujeto mencione una sola para que se le dé el punto correspondiente.

Puntuación máxima: 32

Preguntas	Respuestas Aceptables
1. ¿Para qué tienen alas los pajaros?	Para poder volar.
2. ¿Qué forma tiene una pelota?	Redonda... Circular
3. ¿De qué color es la yerba?	Verde
4. ¿De qué planta se saca el azúcar?	Caña de azúcar ... Remolacha... Betarraga
5. ¿Cuántos meses hay en un año?	12

Preguntas	Respuestas Aceptables
6. ¿Qué es un termómetro?	Instrumento (cosa, etc.) para medir la temperatura... Coger la temperatura... Medir el calor, fiebre.
7. ¿Cuáles son tres idiomas de uso muy frecuente en Latinoamérica?	Español, inglés, portugués, francés, italiano. (Alternativa: Si el sujeto menciona solamente dos idiomas, acredítele un punto).
8. ¿Por qué no se puede caminar sobre el agua?	Porque uno se hunde... Diferencias en la gravedad específica... Tensión insuficiente de la superficie del agua.
9. ¿Qué hay que hacer para contener la sangre de una herida profunda?	Aplicar presión... Poner un torniquete... Ligarla... Amarrarla... Atarla.
10. ¿De qué metal se hacen comúnmente los barcos?	Acero... Hierro... Hierro combinado.
11. ¿De dónde se saca la goma, o sea, el caucho?	Arboles... Savia de los árboles... Un palo... Planta de goma... Petróleo... Alcohol.
12. Rubén Darío fue un hombre famoso. ¿Qué era él?	Poeta... Escritor... Autor... Literato... Novelista.
13. ¿Qué hace el corazón?	Bombea la sangre... Organo de la circulación de la sangre... Parte del sistema circulatorio... Hace circular la sangre.
14. Isabel La Católica fue una mujer famosa. ¿Qué era ella?	Reina de España... Una reina... Ayudó a Colón en su viaje al Nuevo Mundo... Esposa de Fernando.
15. ¿Por qué si se tira un objeto hacia arriba siempre cae?	La fuerza de gravedad... La atracción que ejerce la tierra sobre los objetos.

Preguntas	Respuestas Aceptables
16. ¿Qué es el Vaticano?	La residencia del Papa... Sede del gobierno del Papa... Sede de la Iglesia Católica... Donde vive el Papa.
17. ¿Cuál es la capital de Italia?	Roma
18. ¿En qué dirección viajaría usted si fuera de Río de Janeiro hacia Washington, D. C.?	N... NE... NO
19. ¿Por qué se prende un fósforo si se frota sobre la superficie áspera de la caja de fósforos?	Se calienta... Fricción ... Sube la temperatura... Contiene cierto material químico que al rozarse sale la flama.
20. ¿Cuál es el país que más café produce en el mundo?	Brasil
21. ¿Quién escribió El Quijote?	Cervantes... El Manco de Lepanto... Miguel Saavedra... Saavedra.
22. ¿Cuántas semanas tiene un año?	52
23. Nombre cuatro países de la América del Sur.	Cualesquiera cuatro países situados en la América del Sur: Brasil, Perú, Chile, Argentina. (Alternativa: Si el sujeto menciona solamente tres países, acredítele un punto).
24. ¿Cuándo se celebra el descubrimiento de América?	12 de octubre
25. ¿Por qué se ve el relámpago antes de oirse el trueno?	La luz viaja más rápido que el sonido... La onda de luz es más rápida que la del sonido.

Preguntas	Respuestas Aceptables
26. ¿Cómo es que la levadura hace crecer la masa?	Se fermenta, produciendo burbujas de bióxido de carbono lo cual hace que la masa crezca o suba... Se forman gases... Se fermenta... Se expande... Forma burbujas de aire... Oxidación ... Proceso químico.
27. ¿Por qué la ropa de color oscuro es más calurosa que la de color claro?	Las ropas oscuras absorben el calor del sol... Atraen más el calor... Las ropas de color claro reflejan (repelen) el calor del sol... Retienen más calor.
28. Nombre tres clases de vasos sanguíneos del cuerpo humano.	Arterias, venas, capilares. Los nombres de vasos sanguíneos específicos no son aceptables, pero dé crédito por arteriolas y venillas. (Alternativa: Si el sujeto menciona dos vasos sanguíneos correctos, acredítele un punto).
29. Nombre las tres grandes antiguas civilizaciones indias de Latinoamérica.	Azteca, Maya e Inca. (Alternativa: Si el sujeto menciona dos civilizaciones correctas, acredítele un punto).
30. ¿Quién escribió La Iliada?	Homero
31. ¿Qué es el Corán?	Libro sagrado de los mahometanos... Biblia de los mahometanos. (Se acepta musulmán por mahometanos)... Libro religioso.
32. ¿Dónde está Egipto?	Africa. (Si el sujeto dá alguna otra contestación correcta, tal como "En costa sureste del Mediterráneo", pregunte: "¿En qué continente está Egipto?").

COMPRENSION

INSTRUCCIONES Comience siempre con la pregunta número 1. El examinador puede repetir la pregunta si el sujeto tiene dificultad en recordarla o si el sujeto no responde luego de haber transcurrido de 10 a 15 segundos. En ambos casos, el examinador puede repetir la pregunta pero no debe alterarla ni abreviarla. Si fuese necesario estimular aún más al sujeto, limítese a decir: "Dígame". Si una respuesta no está clara, añada: "¿Quiere explicarme un poco más?" o "Dígame algo más sobre eso".

SUSPENDA la subprueba luego de 5 fallas consecutivas (o sea, 5 respuestas con puntuación de 0).

VALORACION Las preguntas de la 1 a la 15 inclusive reciben cada una puntuación de 2, 1 ó 0. Refiérase a los criterios específicos de valoración y los ejemplos de respuestas en las páginas 71 -80.

Puntuación máxima: 30

Preguntas

1. ¿Por qué lavamos la ropa?

2. ¿Por qué debemos cocinar los alimentos ?

3. ¿Qué debe hacer usted si encuentra en la calle un sobre cerrado, con la dirección puesta y un sello (estampilla) nuevo?

4. ¿Por qué debemos ser muy considerados con los ancianos ?

5. ¿Qué haría usted si está en el cine y es la primera persona que ve humo y fuego?

6. ¿Qué efecto tendría en la salud física de un niño el trabajar muchas horas todos los días cavando zanjas?

7. ¿Por qué cuesta más el terreno en la ciudad que en el campo?

8. ¿Qué quiere decir este refrán: "Aprovecha mientras las brasas están prendidas"?

9. ¿Por qué se deben pagar impuestos?

10. ¿Qué quiere decir este refrán: "Perro que ladra no muerde"?

11. ¿Por qué necesitan gobierno los países?

12. ¿Qué quiere decir la siguiente declaración: "Todos los hombres han sido creados iguales"?

13. ¿Por qué las personas que nacen sordas casi nunca pueden hablar?

14. Si usted se perdiera en el bosque durante el día, ¿cómo trataría de salir?

15. ¿Qué quiere decir este refrán: "Una golondrina no hace verano"?

ARITMETICA

INSTRUCCIONES Al presentar esta subprueba no mencione que consiste de problemas de aritmética. Algunos sujetos tienen una actitud tan negativa o derrotista hacia su habilidad para resolver problemas de aritmética que una presentación de esta índole podría afectar su rendimiento desfavorablemente. Comience la subprueba diciendo solamente: "Ahora vamos a tratar de otra cosa".

Comience siempre con el problema número 3. Si falla en ambos problemas, 3 y 4, administre los problemas I y 2. Si el sujeto falla en ambos problemas, I y 2, suspenda la subprueba, pero si aprueba cualquiera de los problemas I ó 2, prosiga con el número 5.

El examinador puede repetir el problema si el sujeto así lo pide, o si le parece que el sujeto no lo ha comprendido, pero no debe, bajo ninguna circunstancia, alterar la fraseología del problema.

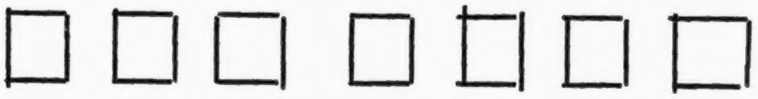
El límite de tiempo se indica en cada problema. Comience a tomar el tiempo al terminar de leer el problema. Si hay que repetir el ítem, debe contarse el tiempo desde que termina la primera presentación. En la Hoja de anotaciones se destina un espacio para anotar el tiempo que el sujeto emplea en la solución de cada problema.

SUSPENDA la subprueba luego de 4 fallas consecutivas.

VALORACION Cada contestación correcta dentro del límite de tiempo recibe una puntuación de I. Acredítele 2 puntos a los sujetos a quienes no se les administren los problemas I y 2. Si la cantidad numérica de la respuesta está correcta se dá crédito por el problema, aun cuando las unidades tales como kilómetros y centavos no hayan sido especificadas. También se dá crédito si el sujeto corrige espontáneamente una respuesta dentro del límite de tiempo.

Puntuación máxima: 16

PROBLEMAS Y RESPUESTAS

Problemas	Tiempo Límite	Respuestas
1. Use los cubos* de la prueba Diseños con Cubos. Coloque 7 cubos con la cara roja hacia arriba, en dos grupos, uno de 3 y otro de 4, o sea  dejando aproximadamente 1/2 pulgada (1.25** centímetros) entre los cubos y 2 pulgadas (5 centímetros) entre los grupos. Pregunte entonces: "¿Cuántos cubos hay en total?"	15"	7
2. Usando el mismo arreglo del problema número 1, pregunte: "Si quita tres cubos, ¿cuántos le quedan?"	15"	4
3. ¿Cuánto es cuatro más cinco?	15"	9
4. ¿Cuántos zapatos hay en dos pares de zapatos?	15"	4
5. Si un hombre compra seis centavos*** en sellos y le dá diez centavos *** al empleado, ¿cuánto debe recibir de vuelta?	15"	4
6. Si en la huerta de su casa usted tiene diez melones y le dá tres a su hermano y cuatro a su padre, ¿cuántos melones le quedan?	15"	3
7. Un cartero camina 25 kilómetros por día durante 6 días, ¿cuántos kilómetros camina en total?	30"	150

---

*Use la palabra bloques si ésta es la de uso común en su país.

** Donde fuese necesario, sustituya el punto decimal por coma.

*** Use la palabra correspondiente en su país.

Problemas	Límite Tiempo	Respuestas
8. Si un hombre compra seis sellos* de dos centavos* y le dá 50 centavos* al empleado, ¿cuánto recibirá de vuelta?	30"	38
9. Si un hombre tiene que caminar 18 kilómetros y ha caminado 7 kilómetros y medio, ¿cuántos kilómetros le quedan por caminar?	30"	10-1/2
10. ¿Cuántas horas tardará un hombre en caminar 40 kilómetros, a razón de 5 kilómetros por hora?	30"	8
11. ¿Cuántas naranjas pueden comprarse con 84 centavos * si una naranja cuesta 7 centavos*?	30"	12
12. ¿Cuántas horas hay en un día y un cuarto de día?	30"	30
13. Un trabajador tarda 16 minutos en llenar dos cajas de melones. ¿Cuántos minutos tardáría en llenar una docena de cajas?	60"	96 minutos 1 hr., 36 min.
14. Un hombre caminó 2.3 **kilómetros y se sentó a descansar. Luego caminó 1.8** kilómetros, más, ¿cuántos kilómetros caminó en total?	60"	4.1
15. Un tanque de agua contiene 400 litros. Esto es 2/3 partes del agua que puede contener si está lleno. ¿Cuántos litros tendrá si se llena completamente?	60"	600
16. Un agricultor tiene una finca* de 60 cuerdas. Si 20% del terreno no se puede cultivar, ¿cuántas cuerdas se pueden cultivar?	60"	48

---

* Use la palabra correspondiente en su país.

**Donde fuese necesario, sustituya el punto decimal por coma.

INSTRUCCIONES Comience siempre con la pregunta número 1. Presente la pregunta diciendo: "Voy a nombrarle dos cosas que se parecen o se relacionan en algo. Por ejemplo, ¿en qué se parecen un lápiz y una pluma?" Si el sujeto contesta que ambos son instrumentos para escribir o implementos para escribir, diga: "Bien", y prosiga con el ítem número 2. Si el sujeto da una respuesta con puntuación de 1 acredítele el punto y diga: "Y se parecen en que son dos instrumentos o implementos para escribir." Si el sujeto menciona una diferencia, no responde, o dice que no se parecen, déle 0 en el item y diga: "Se parecen en que son dos instrumentos para escribir, los dos escriben, los dos tienen punta" y prosiga con el ítem número 2.

Pregunte: "Ahora, ¿en qué se parecen un revólver y una escopeta?" No dé ninguna clase de ayuda en esta pregunta ni en las sucesivas. Presente siempre las preguntas de la misma manera. Omita la palabra "ahora" después de la tercera pregunta.

SUSPENDA la subprueba luego de 5 fallas consecutivas (o sea, 5 respuestas con puntuación de 0).

VALORACION Cada pregunta recibe una puntuación de 2, 1 ó 0. Refiérase a los criterios específicos de valoración y los ejemplos de respuestas en las páginas 81 a la 85.

Puntuación máxima: 30

Preguntas

1. un lápiz	una pluma
2. un revólver	una escopeta
3. un cordón	una soga
4. un baúl	una maleta
5. un hacha	un serrucho
6. un perro	un toro

Preguntas

7.	el oro	la plata
8.	un avión	un caballo
9.	un pantalón	un zapato
10.	el aire	el agua
11.	un ojo	un oído
12.	el norte	el oeste
13.	la sangre	la savia
14.	una poesía	una estatua
15.	un día	un litro

REPETICION DE DIGITOS

Esta subprueba consiste de dos partes, Repetición Directa de Dígitos y Repetición Inversa de Dígitos que se administran separadamente. En cada parte presente los dígitos a razón de uno por segundo, sin agruparlos. Baje el tono de la voz en el último dígito de cada serie. La columna Series indica el número de dígitos en cada ítem.

REPETICION DIRECTA DE DIGITOS

INSTRUCCIONES Comience siempre con el Ensayo I de la Serie 3, diciendo: "Voy a decir unos números. Escuche cuidadosamente y cuando yo termine repítalos usted en el mismo orden en que se los diga".

En cualquier serie, si el sujeto repite correctamente el Ensayo I, prosiga con la siguiente serie. Si el sujeto falla en el Ensayo I, administre entonces el Ensayo II de esa misma serie. Si lo aprueba, entonces prosiga con la siguiente serie. El segundo ensayo de una serie se administra solamente si el sujeto falla el primero. En caso que el sujeto pida que se le repita la serie dígale: "Vamos a tratar esta otra."

SUSPENDA la parte de Repetición Directa de Dígitos cuando el sujeto haya fallado los dos ensayos de una misma serie.

VALORACION La puntuación es igual al número de dígitos de la serie más larga que el sujeto haya repetido sin error bien sea en el Ensayo I o en el Ensayo II.

Puntuación máxima: 9

REPETICION DIRECTA DE DIGITOS

Series	Ensayo I	Ensayo II
(3)	5-8-2	6-9-4
(4)	6-4-3-9	7-2-8-6
(5)	4-2-7-3-1	7-5-8-3-6
(6)	6-1-9-4-7-3	3-9-2-4-8-7
(7)	5-9-1-7-4-2-8	4-1-7-9-3-8-6
(8)	5-8-1-9-2-6-4-7	3-8-2-9-5-1-7-4
(9)	2-7-5-8-6-2-5-8-4	7-1-3-9-4-2-5-6-8

REPETICION INVERSA DE DIGITOS

INSTRUCCIONES Comience siempre la subprueba diciendo: "Ahora voy a decirle algunos números más, pero esta vez cuando yo termine quiero que usted los diga de atrás para adelante. Por ejemplo, si yo digo 7-1-9, ¿qué diría usted?"

Si el sujeto responde correctamente, diga: "Voy a decirle algunos otros" y prosiga la prueba empezando con el Ensayo I de la Serie de 3 dígitos.

Si el sujeto no responde correctamente o no entiende, déle la contestación correcta: "Es 9-1-7". Preséntele otro ejemplo, diciendo: "Recuerde que usted tiene que repetir los números al reves de como yo los diga. Vamos a tratar otra vez: 3-4-8". Si en esta ocasión el sujeto tiene éxito, prosiga la prueba empezando con el Ensayo I de la Serie de 3 dígitos. Sin embargo, si falla el segundo ejemplo, prosiga esta parte usando el Ensayo I de la Serie de 2 dígitos. Si el sujeto responde correctamente a un ejemplo, pero falla en los dos Ensayos de la Serie 3, adminístrele la Serie de 2 dígitos y luego suspenda esta parte. En caso que el sujeto pida que se le repita la serie dígale: "Vamos a tratar esta otra".

SUSPENDA esta parte cuando el sujeto haya fallado ambos Ensayos de la misma serie.

VALORACION La puntuación es igual al número de dígitos de la serie más larga que el sujeto haya repetido a la inversa sin error bien sea en el Ensayo I o en el Ensayo II.

Puntuación máxima: 8

REPETICION INVERSA DE DIGITOS

Series	Ensayo I	Ensayo II
(2)	2-4	5-8
(3)	6-2-9	4-1-5
(4)	3-2-7-9	4-9-6-8
(5)	1-5-2-8-6	6-1-8-4-3
(6)	5-3-9-4-1-8	7-2-4-8-5-6
(7)	8-1-2-9-3-6-5	4-7-3-9-1-2-8
(8)	9-4-3-7-6-2-5-8	7-2-8-1-9-6-5-3

VALORACION TOTAL PARA LA PRUEBA DE REPETICION DE DIGITOS:
La suma total de las puntuaciones obtenidas en Repetición Directa de Dígitos y Repetición Inversa de Dígitos.

Puntuación máxima: 17

INSTRUCCIONES Coloque la tarjeta con la lista de palabras frente al sujeto diciendo: "Deseo que me diga el significado, o sea, qué quieren decir las siguientes palabras. Vamos a empezar con finalizar. ¿Qué significa finalizar?" Al pronunciar la palabra, señálela en la tarjeta que está frente al sujeto. Deberá repetirse la palabra dos veces hasta que el examinador esté seguro de que el sujeto lee la tarjeta simultáneamente, en cuyo caso deberá omitirse la segunda lectura. Proceda entonces con la subprueba hasta que se cumpla con los requisitos de los criterios de suspensión.

Si las respuestas del sujeto a las palabras de la 4-7 ó la 8 son correctas, se puede prescindir de la pregunta formal ("¿Qué significa ________?") y de señalar en la tarjeta. Sin embargo, pronuncie las palabras con corrección y claridad y asegúrese de que el sujeto ha localizado cada palabra en la lista.

Ocasionalmente se hace difícil determinar a través de las respuestas si el sujeto sabe o no el significado de una palabra. En tales casos, el examinador puede decir: "Dígame algo más", o "¿Qué más puede añadir?", o "Explíquese más".

SUSPENDA la subprueba después que el sujeto falle en cualesquiera 7 palabras numeradas consecutivamente, o sea, 7 respuestas con puntuación de 0.

VALORACION Cada una de las palabras recibe una puntuación de 2, 1 ó 0. Refiérase a los criterios específicos de valoración y ejemplos de respuestas en las páginas 86-102.

Puntuación máxima: 80

LISTA DE PALABRAS

1. finalizar
2. gato
3. apresurar
4. comenzar
5. barco
6. escuela
7. mesa
8. ocultar
9. invierno
10. sombrero
11. reunir
12. regocijo
13. encerrar
14. desayuno
15. doméstico
16. admitir
17. meditar
18. caverna
19. transmitir
20. incomparable
21. comer
22. eterno
23. firmeza
24. consumir
25. remontar
26. indispensable
27. remordimiento
28. designar
29. ermita
30. obstruir
31. inventar
32. abolir
33. palpar
34. regularizar
35. fugaz
36. espléndido
37. osado
38. insensato
39. rebasar
40. séquito

DIGITO-SIMBOLO

INSTRUCCIONES El material para esta subprueba se encuentra en la Hoja de anotaciones. Coloque la subprueba[1] frente al sujeto y señalando la clave, diga: "Fíjese en estos cuadros. Note que cada cuadro tiene un número en la parte de arriba y una marca en la parte de abajo. Cada número tiene una marca diferente. Ahora fíjese aquí, (señalando los ejercicios de práctica) los cuadros de arriba tienen números, pero los cuadros de abajo no tienen marca. Usted debe poner en cada uno de estos cuadros la marca que debe ir ahí, de esta forma. (Señale la clave y luego los ejercicios de práctica). Aquí hay un 2, de manera que usted debe poner esta marca. Aquí hay un 1, de manera que usted debe poner esta marca. Aquí hay un 3, de manera que debe poner esta marca".

Escriba en los cuadros correspondientes los primeros tres símbolos como demostración. Luego déle un lápiz al sujeto para que complete los doce cuadros restantes de los ejercicios de práctica. Señale la línea que separa los ejercicios de práctica de la subprueba en sí, y diga: "Ahora haga usted lo mismo con estos números, hasta esta línea."

Si el sujeto aún no puede seguir las instrucciones, repítalas nuevamente demostrando con los próximos tres cuadros y deje que el sujeto complete los cuadros de los ejercicios de práctica que faltan. Luego de la demostración y la práctica, señale el primer cuadro después de los ejercicios de práctica y diga: "Ahora cuando yo le diga que empiece, comience aquí y llene tantos cuadros como pueda sin dejar ninguno hasta que le avise. Listo. Empiece".

Si un sujeto comenzara a omitir cuadros o a hacer un solo tipo de marca dígale: "Háganlos en orden y no deje ninguno en blanco".

LIMITE DE TIEMPO 90 segundos. Comience a tomar el tiempo en el momento de decirle al sujeto que empiece. Recuerde que es muy importante medir el tiempo con precisión.

---

[1] Instrucciones especiales para sujetos zurdos. Coloque frente al sujeto la clave especialmente preparada para estos casos de manera que se vea, ya que el sujeto zurdo a menudo cubre la clave con su mano al escribir en la Hoja de anotaciones.

VALORACION Déle un punto a cada cuadro que tenga el símbolo correcto. Déle medio punto por cualquier símbolo invertido. Los quince ejercicios de práctica no se incluyen en la puntuación.

Puntuación máxima: 110

DIBUJOS PARA COMPLETAR

INSTRUCCIONES Comience siempre con la Tarjeta 1 y presente todos los dibujos a cada sujeto hasta cumplir con los requisitos de los criterios de suspensión. Antes de presentar el primer dibujo, diga: "Le voy a mostrar unos dibujos a los cuales les falta una parte importante. Mire cada dibujo y dígame qué le falta."

Muéstrele al sujeto el primer dibujo y diga: "Ahora, mire este dibujo, ¿qué parte importante le falta?" Limite la presentación de cada dibujo a no más de 20 segundos.

Si el sujeto da la respuesta correcta, prosiga con los dibujos restantes. Al presentar cada uno de los primeros seis o siete dibujos, diga: "Ahora, ¿qué parte le falta a este dibujo?" Si el examinador está seguro de que el sujeto ha captado la idea correctamente, se puede acortar la presentación diciendo: "Y a éste, ¿qué le falta?"

Si el sujeto fracasa en el primer dibujo, diga: "Le falta el ala." Si también falla en el segundo dibujo diga: "Le falta el ojo". A partir del tercer dibujo no preste más ayuda.

Algunas veces un sujeto dice que le falta al dibujo una parte que no es importante. La primera vez que esto suceda se permite decir: "Sí, pero ¿cuál es la parte más importante que falta?" No repita este comentario para ninguno de los dibujos restantes. No se requiere que el sujeto dé el nombre exacto de la parte que falta para recibir crédito, siempre y cuando esté claro que él describe la parte correcta. A veces un sujeto no podrá dar una respuesta verbal, pero señalará una parte del dibujo. Si el examinador está seguro de que el sujeto está señalando correctamente la parte que falta, le dará crédito por ella. Suspenda la prueba después de nueve fallas consecutivas.

LIMITE DE TIEMPO 20 segundos para cada dibujo.

VALORACION Cada respuesta correcta recibe una puntuación de 1.

Puntuación máxima: 25

DIBUJOS PARA COMPLETAR

Item	Parte que Falta
1. Ave	Ala
2. Mujer	Ojo
3. Señorita	Pantalla
4. Cerdo	Rabo
5. Niña	Zapato
6. Candado	Abertura para Llave
7. Niño	Dientes
8. Coche	Rueda
9. Tambor	Palillo
10. Hornilla	Fuego
11. Guitarra	Clavija
12. Perfil de Mujer	Nariz
13. Zapatos	Lengueta
14. Niño con Volantín	Media
15. Hombre con Espejuelos	Puente de los Espejuelos
16. Granero	Portón
17. Bicicleta	Cadena
18. Espejo	Imagen del Brazo
19. Hombre	Dedo
20. Reloj	Uno del Diez
21. Bote	Horquilla
22. Arco y Flecha	Plumas
23. Sol	Sombra
24. Barco	Chimenea
25. Mujer	Ceja

DISEÑOS CON CUBOS

INSTRUCCIONES Comience siempre con el Diseño A. Tome cuatro cubos y diga: "Vea estos cubos[1], todos son iguales. Algunos lados son rojos, otros lados son blancos, y otros, mitad rojo y mitad blanco." Voltee los cubos[1] para mostrar los diferentes lados. Luego diga: Los voy a juntar para formar un diseño. Fíjese". Arregle los cuatro cubos[1] lentamente, formando el diseño de la Tarjeta A pero sin dejar ver la tarjeta al sujeto. Luego dejando el modelo intacto déle otros cuatro cubos[1] al sujeto, señale y diga: "Ahora haga uno exactamente igual a éste". Si el sujeto completa el diseño correctamente, prosiga con el Diseño I. Si el sujeto no logra completar el diseño dentro del límite de tiempo, o si arregla los cubos[1] incorrectamente, recoja los cubos[1] del sujeto, dejando intacto el modelo del examinador, y diga: "Fíjese otra vez en lo que yo hago". Demuestre por segunda vez, utilizando los cubos[1] del sujeto, luego mézclelos y dejando aún intacto el modelo del examinador, diga: "Ahora pruebe usted y asegúrese de que lo hace exactamente igual al mío." Tenga o no tenga éxito el sujeto en este ensayo, prosiga con el Diseño I.

De vez en cuando un sujeto tratará de reproducir exactamente el modelo del examinador, incluyendo el patrón de colores de los lados. Cuando esto ocurra en el Diseño A, o en el Diseño I, dígale al sujeto que sólo es necesario reproducir la parte de arriba de los cubos.[1]

Diseño I. Retire los cubos[1] que sirvieron de modelo para el Diseño A y coloque en su lugar la tarjeta marcada "I". Diga: "Esta vez vamos a colocar los cubos[1] juntos para hacer un diseño igual al de esta tarjeta. Fíjese primero en lo que yo hago." Construya el diseño lentamente, dándole al sujeto la oportunidad de ver que se reproduce el diseño de la Tarjeta I. Entonces mezcle los cubos[1] utilizados en la demostración, colóquelos frente al sujeto, y diga: Ahora fíjese en el diseño y haga uno exactamente igual con estos cubos[1]. Empiece y dígame cuando haya terminado".

---

[1] Use la palabra bloques si ésta es de uso común en su país.

Si el sujeto tiene éxito en este ensayo, prosiga con el Diseño 2. Si falla[1], demuestre por segunda vez, diciendo: "Fíjese otra vez en lo que yo hago. Luego de hacer el diseño, mezcle los cubos[2] y diga: "Ahora trate de hacerlo usted". Tenga o no tenga éxito el sujeto en el segundo ensayo, prosiga con el Diseño 2.

Diseño 2. Mezcle los cubos[2] utilizados en la demostración del Diseño I, coloque frente al sujeto la tarjeta marcada "2" y diga: "Esta vez vez vamos a colocar los cubos[2] juntos para hacer un diseño igual al de esta tarjeta. Fíjese primero en lo que yo hago." Construya el diseño, lentamente, dándole al sujeto la oportunidad de ver que se reproduce el diseño de la Tarjeta 2. Entonces mezcle los cubos[2] utilizados en la demostración, colóquelos frente al sujeto y diga: "Ahora, fíjese en el diseño y haga uno exactamente igual con estos cubos[2]. Empiece y dígame cuando haya terminado."

Si el sujeto tiene éxito en este ensayo, prosiga con el Diseño 3. Si falla, demuestre por segunda vez, diciendo: "Fíjese otra vez en lo que yo hago." Luego de hacer el diseño, mezcle los cubos[2] y diga: "Ahora trate de hacerlo usted". Tenga o no tenga éxito el sujeto en el segundo ensayo, prosiga con el Diseño 3.

Diseños 3-10 : Coloque al frente del sujeto la tarjeta con el Diseño 3 y déle cuatro cubos[2]. Diga: "Ahora, haga uno como éste. Empiece, dígame cuando haya terminado". Cuando el sujeto indique que ha terminado, o al terminarse el límite de tiempo, mezcle los cubos[2] y preséntele el Diseño 4 diciendo: "Ahora haga uno como éste.

---

[1] Inversión del diseño (por ejemplo, invertir el patrón de los cubos en el Diseño I), se considera un error. Antes de administrarle el segundo ensayo diga: "Pero fíjese, va así." Gire los cubos del sujeto en conformidad con el modelo. Se le pide entonces al sujeto que haga el diseño otra vez. Si en el segundo ensayo persiste en la inversión, se considera que el sujeto ha fallado, pero si corrige la inversión en el segundo ensayo, se le dá una puntuación de 2. El mismo procedimiento se sigue con el Diseño 2, pero no con los Diseños 3-10, donde tales inversiones se deben valorar como fallas.

[2] Use la palabra bloques si ésta es de uso común en su país.

Empiece, dígame cuando haya terminado." Siga este procedimiento para todos los diseños subsiguientes.

Al llegar al Diseño 8, saque los otros cinco cubos[1] y diga: "Ahora haga uno como éste, usando todos estos cubos[1]. No se olvide de decirme cuando haya terminado."

LIMITE DE TIEMPO:	Diseños A -2	60 segundos (medidos por separado para cada ensayo en cada diseño)
	Diseños 3-7	60 segundos para cada diseño
	Diseños 8 -10	120 segundos para cada diseño

Si el sujeto reproduce cada diseño correctamente dentro del límite de tiempo, anote el tiempo que se toma para completar cada uno. Para los Diseños 8-10 se dan puntos adicionales por la rapidez en la reproducción del diseño.

SUSPENDA la subprueba después de 4 fallas consecutivas sin tomar en consideración las fallas en ambos ensayos del Diseño A. Las fallas en ambos ensayos del Diseño 1 se consideran como una sola falla. El mismo criterio se aplica para el Diseño 2.

VALORACION: Diseños 1-2: primer ensayo, 4 puntos cada uno; segundo ensayo, 2 puntos cada uno. Diseños 3-7; 4 puntos cada uno. Diseños 8-10: La tabla que se presenta a continuación indica las puntuaciones que se acreditan a cada uno de estos diseños de acuerdo al tiempo que tome al sujeto el completar los diseños. Estas puntuaciones se basan en los diseños que el sujeto logre reproducir correctamente dentro de los límites de tiempo establecidos. Los diseños incompletos o parcialmente correctos no reciben puntuación alguna.

---

[1] Use la palabra bloques si ésta es de uso común en su país.

Puntuaciones para los Diseños 8-10 de acuerdo al tiempo empleado en completarlos.

Diseño	Puntos	4	5	6
8	Tiempo	71"-120"	46"-70"	1"-45
9		81"-120"	61"-80"	1"-60"
10		81"-120"	61"-80"	1"-60"

Puntuación máxima: 46

ORDENAMIENTO DE DIBUJOS

INSTRUCCIONES Presente las tres tarjetas del ítem número 1, AVE, en el orden indicado por los n´umeros al dorso de las tarjetas, y colocando las tarjetas de izquierda a derecha con referencia al sujeto, diga: "En estos dibujos se presenta un cuento acerca de un pájaro que está construyendo un nido, pero los dibujos no están en el orden correcto. Póngalos en el orden correcto para que formen un cuento".

Si el sujeto los arregla correctamente dentro del límite de tiempo, déle 4 puntos y proceda con los otros ítems, diciendo: Aquí tengo otro grupo de dibujos para que usted los ponga en el orden correcto. En cada caso están mezclados y usted los tiene que poner en el orden correcto, de manera que formen el mejor cuento".

Si el sujeto falla en el ítem 1, AVE, demuéstrele la solución correcta, diciendo: "El orden de los dibujos debe ser el siguiente: primero, el pájaro está construyendo el nido; el dibujo que sigue muestra los huevos que el pájaro ha puesto, y el último dibujo muestra al pájaro dándole de comer a los pajaritos recién nacidos". Arregle nuevamente los dibujos en el orden indicado por los números al dorso de las tarjetas y diga: "Ahora, póngalos usted en el orden correcto." Déle dos puntos si el sujeto arregla los dibujos en el orden correcto dentro del límite de tiempo. Déle cero si no lo hace. En cualquiera de los dos casos, prosiga con el ítem siguiente.

Presente las tres tarjetas del ítem número 2, CASA, en el orden indicado por los números al dorso de las tarjetas y diga: "Ahora prueba con este grupo. Ponga estos dibujos en el orden correcto."

Si el sujeto arregla las tarjetas correctamente dentro del límite de tiempo, déle 4 puntos y prosiga con los demás ítems de la subprueba. Si falla, demuestre de la misma manera que en el ítem anterior, diciendo: "Estos dibujos representan a un hombre que está construyendo una casa. El primer dibujo lo demuestra cuando empieza a construirla, el próximo muestra la casa parcialmente construída, y el último muestra la casa cuando ya está terminada y la están pintando." Arregle nuevamente los dibujos en el orden indicado por los números al dorso de las tarjetas, y diga: "Ahora póngalos usted en el orden correcto."

Si el sujeto los arregla esta vez correctamente, déle 2 puntos y prosiga con el ítem siguiente . No dé ninguna otra demostración con el resto de los items de la subprueba. Si el sujeto falla ambos ensayos en cada uno de los items 1 y 2, suspenda la subprueba; de lo contrario, administre todos los items y suspenda luego de 4 fallas consecutivas. En los items del 3-10, presente las tarjetas en el orden indicado, diciendo para cada ítem: "Ahora pruebe con este grupo. Ponga estos dibujos en el orden correcto". Si no está seguro de que el sujeto ha terminado de arreglar la serie, diga: "Cuando haya terminado, dígamelo".

LIMITE DE TIEMPO Refiérase a la tabla que aparece en la página 66. Anote el tiempo exacto que el sujeto se tarda en arreglar correctamente las series 7-10, ya que en estas series se dan puntos adicionales por la rapidez con que el sujeto ordene los dibujos.

VALORACION Series 1-2: primer ensayo, 4 puntos cada uno; segundo ensayo, 2 puntos cada uno. Series 3-10: 4 puntos por arreglo correcto de los dibujos dentro del límite de tiempo. Vea en la tabla siguiente los arreglos correctos, las variaciones aceptadas y la puntuación de cada ítem según el tiempo que el sujeto emplee en completarlos.

Puntuación máxima: 48

PUNTUACIONES SEGUN EL TIEMPO EMPLEADO EN ORDENAR LOS DIBUJOS

	Limite de Tiempo	Arreglo	Puntos adicionales por rapidez			
			2	4	5	6
1. Ave 1er ensayo 2° ensayo	60"	WXY WXY	— 1" -60"	1"- 60" --	-- --	-- --
2. Casa 1er ensayo 2° ensayo	60" 60"	PAT PAT	-- 1" -60"	1"- 60" --	-- --	-- --
3. Jira	60"	EFGH	--	1" - 60"	--	--
4. Caída	60"	SUBE	--	1" - 60"	--	--
5. Asalto	60"	ABCD	--	1" - 60"	--	--
6. Otra	60"	OTRA	--	1"- 60"	--	--
7. Serenata	120"	CANTO	--	26"-120"	16"-25"	1"- 15"
8. Cazador	120"	FUSILA	--	26"-120"	16"-25"	1"- 15"
9. Pibe	120"	LUCHA	--	26"-120"	16"-25"	1"- 15"
10. Galanteo	120"	LEISA o EILSA* ELISA	1" -120" --	26"-120"	16"-25	1" - 15"

---

* LEISA o EILSA - Siempre reciben una puntuación de 2. No reciben crédito adicional por tiempo.

INSTRUCCIONES Presente siempre las figuras en el siguiente orden: el MANIQUI, el PERFIL, la MANO y el ELEFANTE.

Coloque las piezas de cada figura según se indica en los diagramas que aparecen en las páginas 69-70 ocultándoselas al sujeto. (La caja de materiales incluye una pantalla en forma de biombo para este propósito). Una vez que haya colocado las piezas del Maniquí, quite la pantalla, y diga: "Si estas piezas se juntan correctamente, formarán algo. Empiece, y júntelas lo más pronto que pueda".

Coloque las piezas de las otras figuras en sucesión, de acuerdo con los diagramas correspondientes, diciendo en cada caso: "Junte estas piezas lo más pronto que pueda". No permita que el sujeto vea los dibujos de las figuras ya terminadas que aparecen en las páginas siguientes:

LIMITE DE TIEMPO:	Maniquí y Perfil	120 segundos para cada figura
	Mano y Elefante	180 segundos para cada figura

Comience a tomar el tiempo tan pronto como termine de dar las instrucciones. Anote el tiempo exacto que el sujeto se tarde en juntar correctamente las partes de cada figura, ya que se acreditan puntuaciones adicionales por la rapidez con que el sujeto las complete dentro del límite de tiempo.

VALORACION Déle una puntuación de 2 por cada unión correcta que haga el sujeto dentro del límite de tiempo. (Pequeñas desviaciones, por ejemplo, separación en las uniones de no más de un cuarto de pulgada, o de medio centímetro, son aceptables). Las uniones que tienen que ser valoradas están señaladas por una "X" en los diagramas que aparecen en las páginas 69-70. Dé crédito por cada unión correcta aun cuando aparezca en un arreglo parcial. Por ejemplo, si el sujeto une las dos piezas que forman la oreja del Perfil pero sin colocarlas dentro de la cabeza, se da un punto; pero si se unen estas dos piezas correctamente a la cabeza, se dan 3 puntos. Igualmente, si el sujeto une correctamente dos piezas cualquiera de las que forman el Elefante, se le da un punto, aun cuando éstas no estén unidas con ninguna otra parte del cuerpo.

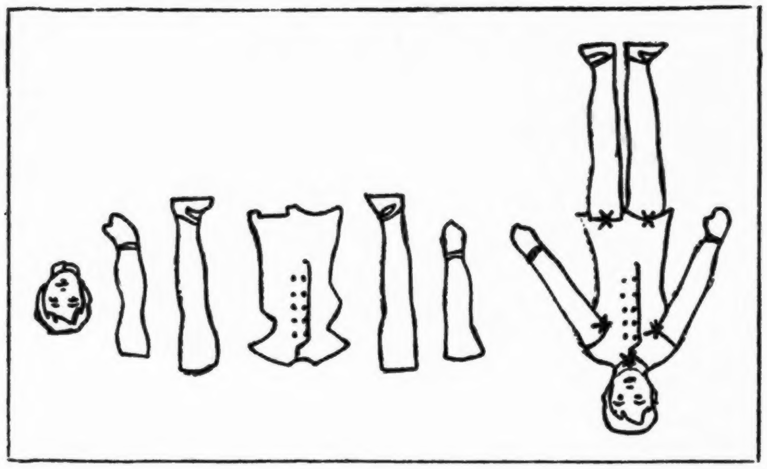
Los puntos que se dan por completar las figuras perfectamente, sin incluir puntos adicionales por rapidez, son los siguientes:

Maniquí	5
Perfil	9
Mano	7
Elefante	8

Debajo de cada diagrama se indican las puntuaciones. Estas incluyen puntos adicionales por rapidez. Tenga en cuenta que los puntos adicionales por rapidez únicamente pueden darse cuando el sujeto haya unido correctamente todas las partes en la figura.

Puntuación máxima: 29

SUJETO

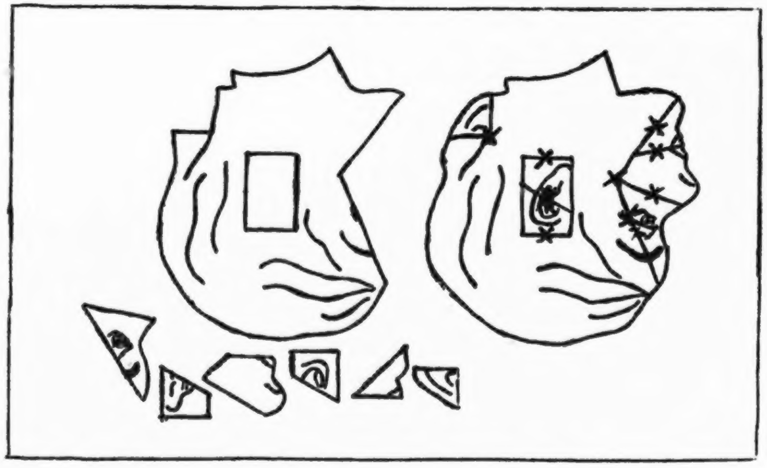


EXAMINADOR

Puntuación por arreglo perfecto del Maniquí según el tiempo empleado en completarlo

5	6	7	8
21"-120"	16"-20"	11"-15"	1"-10"

SUJETO

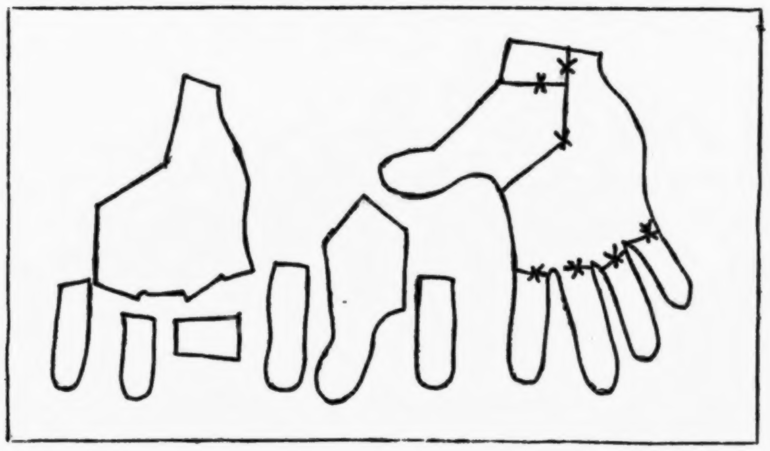


EXAMINADOR

Puntuación por arreglo perfecto del Perfil según el tiempo empleado en completarlo

9	10	11	12	13
46"-120"	--	36"-45"	26"-35"	1"-25"

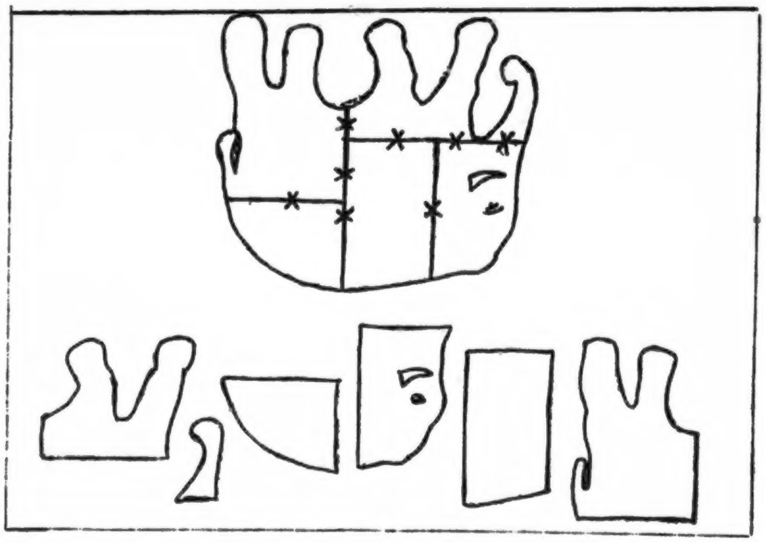
SUJETO



EXAMINADOR

Puntuación por arreglo perfecto de la Mano según el tiempo empleado en completarla.				
7	8	9	10	11
51"-180"	--	41"-50"	31"-40"	1"-30

SUJETO



EXAMINADOR

Puntuación por arreglo perfecto del Elefante según el tiempo empleado en completarlo.				
8	9	10	11	12
51"-180"	--	31"-50"	21"-30"	1"-20"

Criterios generales para la valoración y ejemplos de respuestas: Comprensión

Las respuestas reciben una puntuación de 2, 1 ó 0 dependiendo del grado de generalización y calidad de la respuesta. El examinador tendrá que hacer uso de su buen juicio cuando confronta respuestas poco comunes pues no se han incluído en esta lista todas las posibles contestaciones. No se debe penalizar al sujeto por una expresión pobre. Una respuesta satisfactoria con una puntuación de 2 podría estar mal fraseada. Las respuestas que requieren repregunta indicada por (?) deben ser evaluadas de acuerdo con el criterio establecido.

En cada ítem se presenta el criterio general para cada nivel seguido de algunas contestaciones típicas. La mayor parte de los ejemplos con una puntuación de 0 son típicos de respuestas inapropiadas.

Ejemplos de respuestas

1. ROPA

2 puntos: Cualquier respuesta que incluya, en forma clara y directa, la idea de limpieza.
Porque se ensucia... Para hacerla limpia... Para sacarle el sucio... Para mantenerla limpia...Por la higiene, porque los gérmenes de la ropa sucia afectarían nuestra nuestra salud.

1 punto: Referencia a una idea pobre de limpieza o higiene.
Parte de higiene y de salud... Para sacarle la tierra... Para mantenernos limpios al ponernosla nosotros... Es saludable... Para quitarle las impurezas.

0 puntos: Cualquier respuesta que no haga referencia a la idea de limpieza. Para mejorar la apariencia... Es bueno... Por el uso diario ... Para que no se le pegue el sucio... Para que dure, rendirla...

2. ALIMENTOS

2 puntos: Cualquier contestación que incluya la idea de matar gérmenes, para la salud en general, o para la buena digestión.
Se le quitan las impurezas que tengan... Matar los gérmenes dañinos a la salud... Para hacerlos mas nutritivos...

Para facilitar la digestión... Un alimento crudo es malo para digerirlo.

1 punto: Cualquier respuesta que incluya la idea de que les dá mejor sabor a los alimentos, los hace comestibles, más atractivos y suaves, o una referencia vaga a la salud.
Para no comerlos crudos ... Saben mejor ...Son mejores de comer ...Porque a veces están duros, para ablandarlos...Para adquirir el mejor beneficio de ellos ... Para satisfacer gustos y preferencias del individuo... Para poderse comer ... Para sus atracciones a la vista y al gusto... Es mejor para la salud de uno.

0 puntos: No expresa ninguna de las ideas citadas anteriormente o una idea absurda o errónea.
No podríamos ingerir la mayor parte de los alimentos sin antes cocinarlos... Estamos acostumbrados a cocinarlos... Indice de que estamos civilizados.

3. SOBRE

2 puntos: Cualquier respuesta que se entienda que la carta debe ser depositada en el correo inmediatamente.
Echarla al buzón...Echarla al correo...Dársela al cartero.

1 punto: Reconocimiento de que la carta es propiedad de otra persona pero idea pobre en cuanto a qué hacer con ella.
Averiguaré si se le puede entregar el sobre...Entregarlo a la dirección del remitente... Llevarlo al cuartel de la policía...Llevarlo a persona que dice el sobre.

0 puntos: No tiene idea de qué hacer con la carta o no sabe que la carta es propiedad de otra persona.
Podría quitar el sello y botar la carta...Abrirla y leerla...

4. ANCIANOS

2 puntos: Cualquier respuesta que contenga la idea de respeto por la edad, la idea de la sabiduría que da la experiencia, o disminución de la agilidad física.
Considerarlos y respetarlos por su edad y experiencia... Ellos son mayores que nosotros y debemos tenerlo en consideración... Son dignos de admiración y respeto... Tienen ya experiencia, nos han dado ejemplos ...Tienen menos agilidad física.

1 punto : Idea específica de amor por el vecino, nos ponemos viejos, también limitación física mal expresada, o necesidad de ayuda.
Se merecen mejor trato debido a sus limitaciones ... Son mucho más débiles que nosotros ...Necesitan más nuestra ayuda...Ellos no están a la par de valerse por ellos mismos... Uno algún día llega a esta etapa... La satisfacción que produce dar ayuda a quien la necesita... Por caridad con el prójimo...Porque están incapacitados físicamente muchas veces... Ellos fueron los que nos criaron y nos ayudaron.

0 puntos : Una respuesta vaga o no tiene idea de los conceptos ya citados. Su estado físico se encuentra en decadencia y sus energías son menores que las de nosotros ... Su estado no les permite ver las cosas razonablemente... Para que el anciano no se sienta rechazado... Por el bien de la sociedad.

5. CINE

2 puntos : Reconocer que una persona con autoridad que se encuentre presente debe ser notificada, como el administrador o el acomodador.
Avisarle a la gerencia... Le aviso a uno de los señores que están en la puerta ... Iría al administrador y avisarle... Salgo y se lo digo al encargado.

1 punto : Reconocer que hay que actuar responsablemente y que se debe tomar alguna medida aunque ésta no sea de efectividad inmediata.
No gritar y dar aviso... Llamaría a los bomberos... Conseguir un extinguidor.

0 puntos : Descripción de acciones que crearían un pánico o que no impedirían el desastre.
Saldría inmediatamente...Avisar fuego! fuego! Salgan que se queman! ... Grito y salgo corriendo... Huir... Daría la alarma (?)...Avisar a los demás...Avisar.(?)[1]

---

[1] (?) Haga repregunta siempre que aparezca este signo.

6. NIÑO

2 puntos: Cualquier contestación que contenga la idea de que retarda el desarrollo o de que es detrimente o impide el desarrollo físico adecuado.
Trabajar muchas horas le impediría su desarrollo físico...Retrasa el desarrollo físico normal del niño... No se puede desarrollar adecuadamente, impide su desarrollo... El desarrollo del nene se atrasa y no crece bien.

1 punto: Un efecto específico o idea general de que es nocivo para la salud del niño. Desgaste físico, debilitamiento de energía, adelgaza...Decaimiento, debilidad física... Fatiga, asma, se agota por cansancio... Se le podrían doblar los huesos de la espina dorsal...Le restaría demasiado de muchas energías ...Se enfermaría... Lo fatigaría en extremo... Lo debilitaría, eso afectaría la salud... Se cansa.

0 puntos: No representa peligro para el desarrollo y la salud o idea vaga de alguna relación con la salud.
Un niño no tiene la resistencia física para esa clase de trabajo... No puede hacer nada... Pudiera adquirir fácilmente la poliomielitis... No sé si le haría provecho, a muchos le hace bien, otros se consumen... Pueden coger parásitos.

7. TERRENO EN LA CIUDAD

2 puntos: Mención de demanda con la implicación de limitación en la oferta. En la ciudad hay más demanda y más escasez del terreno que en el campo...Se relacionaría con la ley de manda y oferta, que los solares del pueblo que son los más buscados son los que tienden a coger más valor... Hay que pagar más por la demanda que tiene... Porque hay menos terreno y vive más gente, así sube el precio.

1 punto: La idea de oferta sin implicar demanda, o la mención de más de una facilidad localizada en la ciudad como: teatros, tiendas, transportación, controles de seguridad, utilidades, facilidades sanitarias, calles.
Acá hay muchas más comodidades que en el campo... Hay poco terreno disponible en la ciudad...Localización del

terreno ya que en la ciudad está cerca de negocios e industrias...Porque tienen más y mejores facilidades, las escuelas, transportación, medios de diversión, etc...El terreno está ya urbanizado.

0 puntos: Respuesta que demuestre desconocimiento total o parcial de la ley económica envuelta.

En la ciudad hay más oportunidades y progreso... Hay más personas dispuestas a pagar y más dinero en la ciudad... La vivienda por lo regular es más cara que en el campo... La ciudad es el centro de industrias, comercios... La vida es más fácil en la ciudad... Hay más movimiento para salir y los niños jugar... Por el lujo es más bonito y más vistoso.

8. BRASAS

Si la respuesta es otro proverbio, pregunte el significado y valórelo de acuerdo con el criterio.

2 puntos: Una generalización abstracta.

Que mientras encuentre que tengo facilidades para llevar a cabo algo debo aprovecharlo...Que uno debe aprovechar las situaciones mientras se mantiene vivo el interés hacia ella... Uno debe actuar cuando tiene la oportunidad y no dejarla perder...Aprovechar las oportunidades o la ocasión de actuar en el momento, no dilatar.

1 punto: Un ejemplo específico o relacionado que no sea exactamente una generalización equivalente.

Tratar de sacar lo mejor de una situación mientras está vivo... Aprovechar el tiempo mientras uno es joven y no esperar que uno envejezca sin hacer nada...que se aproveche uno de la oportunidad...Que hay que aprovechar el tiempo...Hacer uso de las situaciones más favorables.

0 puntos: No reconoce que el dicho es un proverbio. Respuestas sobre"brasas", explicación vaga sobre aprovecharse de una situación en el nivel personal, traducción literal o elaboración.

Aprovechar la vida...Sería cocinar algo... No dejar que se apague el fuego.

9. IMPUESTOS

2 puntos : Para sostener los gastos del gobierno y de las instituciones gubernamentales. (Ciudad, comunidad o estado se pueden mencionar en vez de gobierno).
Son una manera o forma de mantenerse el gobierno en cuanto al aspecto económico ...Para que el gobierno lleve a cabo sus actividades que son necesarias para el bienestar de la comunidad...Para ayudar a sostener los gastos públicos que el gobierno tiene...Para poder sostener los programas del gobierno.

I punto: Mención específica de dos o más instituciones mantenidas por el gobierno o bienestar público en general .
Para que el gobierno nos construya carreteras y ayude al bienestar del país...Para el bienestar del público o país, el mejoramiento de la comunidad...Crear bienes públicos como escuelas, parques, etc... Se usan estos impuestos para el fondo público, para hacer carreteras, hospitales, etc.... Porque con ellos es que el gobierno hace las obras públicas... Para pagar sueldos a los empleados del Gobierno.

0 puntos: Referencia a una institución específica, organización o trabajo en donde no demuestre haber captado la idea de que los impuestos son usados para la estructura gubernamental en total.
El estado nos da algo a nosotros y tenemos que dar algo a ellos... Por medio de impuestos el gobierno da beneficio al pueblo... Para saber que los beneficios que recibimos son parte nuestra... El gobierno nos lo exige... El gobierno dice que ese dinero se devolverá en progreso... más progreso (?)

10. PERRO

Si la respuesta es otro proverbio, pregunte el significado y valórelo de acuerdo con el criterio.

2 puntos: Una generalización abstracta.
Que aquel individuo que habla mucho y se alaba mucho por lo regular no hace lo que dice...Se refiere a personas que mucho hablan y al cabo sus palabras no significan nada... "Que no todo lo que aparenta ser es" ... La gente muy bocona es la que menos hace.

1 punto: Un caso específico o una generalización relacionada pero no exactamente equivalente.

"No todo lo que brilla es oro" (?)... Mucho ruido y pocas nueces (?)... Se aplica a las personas que tienen un carácter violento o que exageran su actuación... A veces nos dejamos llevar por alardes; sin embargo, lo que se pregona no es lo que sucede sino otros que se pregonan menos... Para hacer algo no debe estar hablando, sino actuar... El que dice que va a hacer una cosa y no la hace.

0 puntos: No reconoce que la aseveración es un proverbio, o cambia el sentido a un nivel bien específico.

El perro ladra y a veces no muerde; a veces muerde callao... Implica que las personas que no esperan mucho rara vez son personas de acción... En la vida debe hacerle frente a cualquier problema sin temor...Que las personas que son de carácter fuerte por lo general no son más comprensivas... Que no le tengan miedo a una persona porque no va a hacer daño.

II. GOBIERNO

2 puntos: Una aseveración que exprese una razón fundamental, por ejemplo, para crear y regular las leyes que mantienen la sociedad en paz y en orden, o dos o más funciones principales del gobierno.

Para mantener la paz... Dirigir las leyes y mejorar el bienestar de todos...Para garantizar los derechos civiles del pueblo... Para conservar la economía y el orden social de la nación... Para poder dirigir los destinos del pueblo.

1 punto: Una idea específica o una función del gobierno. La idea de conservar el bienestar de la sociedad, o la idea de orden o dirección sin mayor explicación.

Para dirigir el país... Para que se viva mejor...Para justicia del pueblo y buen funcionamiento...Para que haya sociedad... Para tener una autoridad.para regir...Para poder controlar las leyes...Para los pueblos regirse...Para poner el orden...Porque sin gobierno se formaría una desorganización social y provocan líos...Para el bienestar del país.

0 puntos: No tiene idea o tiene una idea incorrecta sobre el gobierno o sus funciones.

Sin gobierno no se puede vivir...Para evitar guerras... Traer responsabilidad y respeto a la nación...Tiene que tener un hombre que represente al país...Para poder gobernar...Para que no haya delincuencia juvenil... Porque todas las personas tratarían de gobernar ... Son mandatorios en cada país.

12. IGUALDAD

2 puntos: La idea de que el hombre como ser humano tiene ciertos derechos inalienables; mención de dos o más de estos derechos, o la idea de que todos los hombres como seres espirituales son iguales ante el Creador.

Todas las personas ante Dios somos iguales...Todos tienen los mismos derechos civiles...Todos los hombres tienen el derecho de vivir, de libertad, de tener propiedad privada.

I punto: Una idea vaga o respuesta parcial de dos puntos sin explicar.

Dios creó a todo el mundo igual (?)...Son hijos de Dios... Que a todos Dios nos hizo de la misma manera con las mismas libertades (?)...No debe haber distinción entre uno y otro...Que todos tenemos el mismo derecho en la vida y la misma obligación, nos debemos considerar unos a otros.

0 puntos: No tiene idea o tiene idea incorrecta del significado de la aseveración.

Que todos son iguales (?)... Todos los hombres son tratados igualmente...Todos nacen iguales (?) ... Que nos debemos llevar todo el mundo igual ...Que nadie es mejor que nadie (?)... Que todos los hombres deben ser iguales... Todos vienen del mismo origen (?)...Que no tenemos ninguno defectos... Que no deben existir diferencias de razas.

13. SORDOS

2 puntos: Reconocer que uno tiene que oír hablar u oír sonidos envueltos en el lenguaje hablado antes de poder repetirlos.

Porque nunca han oído el sonido y el hablar es repetir sonidos que se aprenden por sonidos... La persona sorda no puede imitar sonidos de otros... No oye los sonidos, no se puede repetir o imitar... Porque el lenguaje casi siempre se aprende escuchándolo.

1 punto:	La idea de que oir y hablar están relacionadas sin saber cómo. Nunca han tenido la oportunidad de oír, escuchar, hablar, no conocen los sonidos... Para poder hablar la persona tiene que aprender el sonido externo y así acoplarse al lenguaje en sí ... No pueden captar sonidos.
0 puntos:	No tiene idea sobre la interdependencia entre la audición y el habla. Es una cosa fisiológica de herencia y difícil de arreglar... Se le interrumpe la habilidad de las cuerdas vocales...Porque los órganos de la voz no son desarrollados bien... Es una obra mandada por Dios... Hay relación entre el oído y las cuerdas vocales...Porque no pueden oír bien la frase adecuada según se habla.

14. BOSQUE

2 puntos:	Cualquier explicación del uso de un fenómeno natural para encontrar una manera de salir o un enfoque sistemático al problema. Por medio de la posición del sol trataría de localizar un punto direccional, ejemplo norte, y guiarme por dirección de la sombra de los objetos... Si hay un río cercano, el muzgo crece en la parte norte de los árboles, caminaría en esa dirección...Dirigiéndome por el sol que sube por el este, yo seguiría derecho en esa dirección.
1 punto :	Mención de un método impensado (al azar) para poder salir, o una respuesta parcial de 2 puntos sin explicar. Guiarme por los rayos del sol... Me conduciría por el sol...Mirando u observando el movimiento o la dirección del sol... Caminar en línea recta, primero me fijaría una dirección y luego la seguiria...Esperaría la noche y me guiaría con las estrellas... Seguir un río.
0 puntos:	Uso de fenómenos poco confiables o absurdos, o depender de otras personas. Caminando como una loca tratando de conseguir el lugar...Buscar el sitio por donde entré... Creo seguiría caminando hasta llegar a la carretera o a algún lugar menos poblado de vegetación...Siguiendo siempre la meta opuesta a la que yo venía... Teniendo idea de por dónde entré y caminar ligero antes de que me coja la noche. Pediría auxilio...Viraría hacía atrás.

15. GOLONDRINA

Si la respuesta es otro proverbio, pregunte el significado y valórelo de acuerdo con el criterio.

2 puntos: Una generalización abstracta.

Una papa no hace un caldo – se necesitan varios estudiantes para hacer una clase... Este solo factor no va a determinar todo lo demás... Por algo que suceda no se puede hacer juicio sin tener más evidencia...Una sola cosa no es la que forma el todo de una cosa importante... Que uno no puede generalizar por un sólo hecho.

1 punto: Un caso específico, o una generalización relacionada pero no exactamente equivalente.

Que se vea una golondrina no necesariamente quiere decir que ya el verano ha llegado...Una cosa no es lo que se necesita para hacer algo, sino que una serie de cosas... La vida simplemente con que uno logre un objetivo no logra avanzar el todo.

0 puntos: No reconoce que la aseveración es un proverbio, o cambia el sentido del proverbio a un nivel bien específico.

Que algo que pasa rápidamente no deja una raíz... La golondrina por lo general se manifiesta en la primavera... Emigran, no pasan el verano en sitio específico... El viajar mucho no permite estabilidad. No se puede lograr mucho si se pasa saltando de un lado a otro... Yo diría que la vida no es solamente un término de años, sino que es constante, que no se para...Una sola persona no puede dirigir.

Criterios generales para la valoración y ejemplo de respuestas: Analogías

A continuación se indican los criterios generales que se usan en la valoración de los ítems de esta subprueba:

2 puntos:	Cualquier clasificación general que envuelva un parecido fundamental entre ambos miembros del par.
I punto:	Cualquier propiedad específica común a ambos miembros del par y que constituya una semejanza pertinente.
0 puntos:	Propiedades específicas de cada miembro del par, generalizaciones que no son pertinentes, o diferencias.

Los ejemplos que se incluyen a continuación son respuestas típicas de la población que se usó para la estandarización de la prueba: Se ilustran las puntuaciones asignadas a cada una de éstas. Puede asumirse que las frases "ambas son" o "ambas tienen" preceden a la mayoría de las respuestas.

Ejemplos de respuestas.

I. LAPIZ-PLUMA

2 puntos:	Instrumentos para escribir...Implementos que usa el hombre para escribir.
I punto:	Para escribir...Tienen punta... Los dos escriben...Se usan en la escuela... Se usan en la oficina.
0 puntos:	Los dos tienen el mismo uso...El lápiz usa carbón y la pluma tinta...El lápiz es de palo y la pluma, no.

2. REVOLVER -ESCOPETA

2 puntos:	Armas de fuego.
I punto:	Armas... Se usan en la guerra... Los dos disparan... Los dos están hechos de metal...Los dos matan...Ambos usan balas...Ambos se usan para defendernos...Los dos usan pólvora...Los dos tienen gatillo.
0 puntos:	Los dos hacen explosión.

3. CORDON-SOGA

2 puntos: Son cuerdas...Materiales hechos de fibras entretejidas, generalmente largas y de forma redonda.

1 punto : Para amarrar...Para atar...Para ceñir...Para colgar... Los dos se usan más o menos con el mismo propósito y están hechos más o menos de la misma forma... Son hechos de hilo.

0 puntos: Se parecen mucho; el cordón es finito y la soga...Los dos son largos.

4. BAUL-MALETA

2 puntos: Equipaje... Son valijas... Cofres.

1 punto : Se usan cuando uno viaja... Para echar ropa...Para guardar cosas o ropa... Se trancan, tienen llave.

0 puntos: Son para cargar al hombro.

5. HACHA-SERRUCHO

2 puntos: Son herramientas...Instrumentos de carpintería...Son instrumentos cortantes.

1 punto : Para cortar madera...Instrumentos..Se usan para cortar... Tienen filo... Se construyen de un material igual... Contienen madera y acero...Pican madera...Sirven para trabajar en madera... Son de hierro.

0 puntos: El serrucho de serruchar y el hacha de picar... Sirven para trabajar...Se usan para trabajar.

6. PERRO-TORO

2 puntos : Animales...Animales irracionales...Cuadrúpedos... Mamíferos.

1 punto: Ambos tienen una cola y un rabo...Tienen piel... Los dos comen...Tienen ojos...Cuatro patas...Son de utilidad para el hombre...Tienen vida... Los dos corren.

0 puntos: Perro muerde y el toro faja...Son bravos.

7. ORO-PLATA

2 puntos: Son metales...Son elementos...Son minerales.

I punto : Son valiosos...Son costosos... Son duros... Se usan para hacer prendas... Se usan para hacer sortijas, etc.... Los dos brillan... Se usan como monedas... Se usan para hacer dinero... Se usan para comprar muchas cosas.

0 puntos: Son monedas.

8. AVION-CABALLO

2 puntos: Métodos o medios de transportación...Medio de locomoción.

I punto : Se usan para transportación...Sirven para conducir a una persona...Medios de comunicación... Ambos cubre distancias y transportan pasajeros... Ambos transportan... Ambos tienen caballos de fuerza... Ambos tienen cola... En los dos se montan... Los dos corren.

0 puntos: El avión vuela y el caballo corre...Ambos usan comestibles...Ambos van y vienen...El avión vuela y el caballo vuela por la tierra...Uno corre y otro para viajar.

9. PANTALON -ZAPATO

2 puntos: Artículos de vestir...Son prendas de vestir... Vestuario... Vestimenta.

I punto : Los dos sirven para vestir... Se usan para cubrir parte del cuerpo. Efectos de ropa...Que se ponen para vestir... Son para usarlos... Son ropa...Son de uso personal... Sirven para protección del cuerpo.

0 puntos: Ambos tienen que meter la pierna para llevarlo... Son fabricados...Tienen dos patas... Vestidos.

10. AIRE-AGUA

2 puntos: Medios para la subsistencia... Necesarios para la vida... Ambas fuerzas naturales poderosas.

I punto : Son elementos... Ambos son medios de transportación y de comunicación... Tienen elementos comunes (explicado)... Tienen oxígeno... Sin color... Se mueven... Sirven para mover maquinaria... Están en la atmósfera... Corre el aire y corre el agua... Son necesarios.

0 puntos: Recursos naturales...Dos substancias: una gaseosa y la

líquida...El agua se toma y el aire se respira... Ambos refrescan...Son fríos...Los dos vienen de lo alto...Uno seca, otro moja...Son importantes... Tienen peso y ocupan espacio.

11. OJO-OIDO

2 puntos: Ambos son órganos de los sentidos. Son órganos sensoriales...Dos sentidos...Receptores de estímulos... Para captar sensaciones...Organos de Percepción.

1 punto: Son órganos...Sirven para sensibilidad del individuo, para captar sonidos o imágenes...Son partes sensitivas... Los dos ayudan al entendimiento...Con los dos uno se orienta... Ambos pertenecen a la cara...Los dos producen placer estético... Nos ayudan a comunicar con el medio ambiente...A través de ellos se perciben sensaciones...Miembros del cuerpo... Partes del cuerpo... Los dos se necesitan iguales... Los dos perciben.

0 puntos: Son parte del sistema del hombre...Los dos llevan la misma letra... Son redondos... Indispensables al hombre... El ojo ve y el oído oye... Podemos distinguir con los dos.

12. NORTE-OESTE

2 puntos: Direcciones...Puntos cardinales principales...Polos de dirección...Puntos de referencia para tomar orientación.

1 punto : Ambos son posiciones...Ambos orientan...Dicen a dónde vamos...Ambos son polos...Puntos de localización.

0 puntos: Los dos tienen la letra o... Medidas de dirección... Distancia... Costa sur y costa norte... Tienen direcciones opuestas.

13. SANGRE-SAVIA

2 puntos: Los dos son líquidos vitales en diferentes cuerpos vivientes...La sangre circula en el cuerpo y lleva alimentos, la savia is igual para el árbol ...La sangre es al ser humano como la savia es a las plantas...Da vida en diferentes cuerpos, plantas y seres humanos.

1 punto : Son líquidos...Las dos transportan el alimento al organismo... Ambas corren...Ambas fluyen...Las dos son necesarias... Medios para vivir (hombre y planta).

0 puntos: La sangre es del cuerpo humano y la savia es del animal y vegetal...En el color...Las dos pertenecen al corriente sanguíneo.

14. POESIA - ESTATUA

2 puntos: Expresiones o concepciones artísticas...Obras de arte... Productos del arte.

1 punto : Son manifestaciones de sentimientos... Lo hizo un artista...Manifestaciones de la belleza...Las dos son resultado de inspiraciones...Tienen valor artístico... Creadas por el hombre...Reflejan lo que el autor y el escultor sienten...Simbolizan algo.

0 puntos: La poesía casi siempre está dedicada a una persona y las estatuas a las personas... Se pueden rimar...Ambos tienen autor... Los dos tienen historia... Hechas por el hombre... Las dos describen.

15. DIA-LITRO

2 puntos: Medidas ...Unidades de medición...Hacen una medida fija.

1 punto : Tienen medidas...Pueden servir para medidas... Los dos son unidades... Tienen contenido...Ambos pueden dividirse en cuatro partes iguales.

0 puntos: Los dos son claros...Los dos tienen fin.

Criterios generales de valoración y ejemplos de respuestas: Vocabulario

En términos generales se acepta cualquier significado reconocido de la palabra, sin tomar en cuenta la elegancia en la expresión. No obstante, toda respuesta que denote pobreza de contenido se penaliza en cierta medida. Cualquier contestación que indique que el sujeto sólo tiene un conocimiento vago del significado de la palabra no recibe crédito completo. Todas las respuestas a los items de la subprueba reciben una puntuación de 2, 1 ó 0.

Los siguientes son principios generales para valorar las respuestas de los items de Vocabulario .

2 puntos:

1. Un buen sinónimo
2. Algún uso principal
3. Uno o más rasgos fundamentales, que definan la palabra.
4. La clasificación general a la cual pertenece la palabra.
5. Varias características descriptivas que, tomadas en conjunto, indiquen conocimiento del significado de la palabra, aunque no la definan con precisión.
6. En los verbos se requiere que el sujeto dé ejemplos claros y precisos de acción, o ejemplos que demuestren una relación causal.

I punto:

1. Respuesta que demuestren pobreza de contenido, aunque no sean precisamente incorrectas.
2. Un sinónimo vago o inexacto.
3. Un uso menor, sin elaborar el significado.
4. Atributos correctos, pero que ni definen la palabra ni la diferencian con exactitud.
5. Ejemplos en que se hace uso de la misma palabra, pero sin elaborar su significado.
6. Definiciones correctas de un derivado de la palabra, como por ejemplo: "apresura" en vez de "apresurar", u "obstrucción", en vez de "obstruir".

0 puntos:

1. Respuestas obviamente incorrectas.
2. Uso de verbalismos, como por ejemplo, "reparar un carro", si luego de interrogar al sujeto, éste no demuestra un verdadero conocimiento del significado de la palabra.
3. Respuestas indefinidas o que demuestren marcada pobreza de contenido aún después de la repregunta.

Cada pregunta específica de Vocabulario tiene un criterio general para la puntuación de 2, seguido de varios ejemplos de respuestas. En algunas ocasiones se dan criterios generales para las respuestas con puntuaciones de 1 ó 0, con sus respectivos ejemplos. En otras ocasiones no hay generalización apropiada para establecer puntuaciones de 1 ó 0 aparte de los incluídos en los criterios generales de valoración. En los casos en que esto ocurre, solamente se incluyen ejemplos de respuestas. Desde luego, esta lista de respuestas sólo contiene algunas de las muchas posibles contestaciones que los sujetos pueden dar. Su función es simplemente suplementar las definiciones de diccionario y los criterios generales de valoración, de manera que se facilite la tarea de valorar esta prueba con precisión.

Ejemplos de respuestas.

1. FINALIZAR

2 puntos:	Concluir (una obra). Extinguirse o acabarse una cosa. Terminar. Dar fin a algo... Acabar con algo.
1 punto:	Cuando llega a una conclusión...Cuando se escucha algo y se escribe la última oración... Cosa que tenga que lo vaya a terminar ...Algo que se termina...Finalizar un trabajo...El final .
0 puntos:	Finalizar algo o cosa.

2. GATO

2 puntos:	Mamífero carnicero, doméstico de la familia de los félidos. Máquina compuesta de un engranaje de piñón y cremayera para levantar grandes pesos, de automóvil, etc. Instrumento de hierro para asir y transportar la madera.

Animal doméstico...Un animal...mamifero cuadrúpedo.

1 punto: Hombre astuto, ratero...cuadrúpedo...Tiene 4 patas ... que coge ratones...Para levantar el carro... Un nombre de animal .

0 puntos: Tiene pelo...Camina y Corre... Tener un gato.

3. APRESURAR

2 puntos: Dar prisa (a uno). Hacer ligero, de prisa.
Avanzar...Aligerar...Apurar... Ir rápido... Darse prisa...Ajorar...Avanzar una cosa...

1 punto: Ligero...Rápido...Cuando una va de prisa... Tener prisa...Cuando uno quiere terminar temprano lo que está haciendo...Hacer las cosas más rápidas... Hacer las cosas ligero.

0 puntos: Coger algo y no soltarlo... Ligereza...Apresurar una cosa, correr.

4. COMENZAR

2 puntos: Empezar. Dar o tener principio a una cosa. Iniciar.
Dar comienzo (a cualquier cosa)... Principiar.

1 punto : Algo que se inicia...Cuando al hacer una obra lo primero que hace...Comenzar un trabajo...El principio.

0 puntos: Hacer de nuevo una cosa... La introducción de algo... Comenzar una cosa...

5. BARCO

2 puntos: Construcción de madera o metal dispuesto para flotar, correr o transportar por el agua. Navío.
Medio de transportación por el agua...Usado para navegación...Usado para transportar y viajar por el mar ... Embarcación...Nave...Buque hecho de hierro, madera, que pasa por los mares...Para navegar por el agua.

1 punto : Medio de comunicación marítima... Medio de transportación...Lleva personas de un sitio a otro...Se usa para embarcaciones...Forma de una goleta...Para viajar ...

Para cargar...De embarcar... En barco se puede viajar...Algo de madera que va por el agua... Para llevar pasajeros, carga...Caminar por el mar... Canoa, lancha, vapor, bote...Un vehículo.

0 puntos: Hecho de metal...Medio más antiguo que usó el hombre...Es muy cómodo... Lo que los países usan para comercio o para la defensa nacional ... Embarque

6. ESCUELA

2 puntos: Establecimiento público o privado donde se dá cualquier género de instrucción y especialmente la primaria... Plantel de enseñanza...Centro de enseñanza ... Institución donde se enseña...Aula.

I punto : Donde educan a los niños...Que uno estudia en ella... Donde uno aprende...Donde los niños van a estudiar... Sitio donde se estudia...Le dan clases a uno...Colegio.

0 puntos: Estudio...Enseñanza...Educación...Un salón.

7. MESA

2 puntos: Mueble compuesto por una tabla lisa sostenida por uno o varios pies.. Un mueble... Tabla lisa con cuatro patas.

I punto: Mesa en que se come...Para sentarse a escribir... Donde uno come.

0 puntos : Un objeto de la casa...Quitar o poner la mesa... Lo que se usa para comer...Sentarse a la mesa.

8. OCULTAR

2 puntos: Esconder. Poner en un lugar o sitio retirado o secreto para que no sea encontrado fácilmente. Impedir que sea vista. Encubrir. Poner fuera de la vista. No permitir que otros vean...Guardar algo en secreto, tapar... Reservar...Disimular (sentimiento o pensamiento)

I punto : Mentira, no decir la verdad...Tener oculto...Secreto... Algo que uno vea y lo oculta, no diría nada...Negar una cosa...No divulgar...Escondido...Cuando uno hace algo malo y no quiere que el papá lo sepa.

0 puntos: Guardar algo... No decir nada... Cuando uno quiere que algo se vea... No hacer nada... Tener guardado algo.

9. INVIERNO

2 puntos: Una de las cuatro estaciones del año, la más fría de las estaciones. Estación en que cae nieve en distintos países... Cuarta estación del año en América.

1 punto : Una de las estaciones del año... Tiempo frío...Cuando llega el frío... Cuando está nevando...Nieve...Hace frío...Frío...Epoca, estación o temporada del año... Cuando el sol sale más tarde y oscurece más temprano.

0 puntos: El clima cambia... Temporada en un país...Cuando los árboles pierden las hojas...Cuando pasa el verano, viene el invierno...Lo que decimos cuando tenemos frío.

10. SOMBRERO

2 puntos: Prenda de vestir para cubrir la cabeza que consta de copa y ala... Vestuario que se usa en la cabeza.

1 punto: Gorra...Pamela de mujer...Boina...Capacete...Casco... Pava...Birrete...Para cubrir la cabeza...Algo que uno se pone en la cabeza.

0 puntos: Objeto que el hombre usa para taparse del sol..Que uno usa para salir e ir a trabajar.

11. REUNIR

2 puntos: Volver a unir. Juntar. Congregar. Aunar. Agrupar...Unir un grupo de personas juntas...Unirse... Hacer un conjunto...Ajuntar.

1 punto : Unir...Reunir grupo de personas...Coleccionar...Recolectar... Se refiere a la reunión...Cuando las personas se reúnen para dar una asamblea... Recoger...Una reunión... Reunir dinero...Convocar, colectar, acumular, aglomerar, recopilar, confederar... Algo que unimos.

0 puntos: Organizar un grupo... Conservar o guardar un grupo o algo... Un grupo de algo o alguien... Hablar con un grupo de personas en un lugar... Grupo... Guardar dinero... Economizar.

12. REGOCIJO

2 puntos: Júbilo...Alegría...Alborozo...Contento.
I punto : Gozo...Felicidad...Algo agradable...Sentir un gran placer por algo...Estar alegre o contento... Lo que sentimos cuando obtenemos éxito... Una persona que está alegre...Satisfacción...Uno se regocija si viene cualquier persona a la casa que hace tiempo uno no ha visto... Alegre, alegrarse...
0 puntos: Paz, tranquilidad...Sentirse orgulloso de algo... Es ser bueno...Halago... Algo que ha sido de nuestro agrado... Esmero...Entusiasmo... Se siente bien.

13. ENCERRAR

2 puntos: Meter a una persona o cosa en parte donde no puede salir, incluir (en el sentido figurado). Contener.
Poner en un sitio para que nadie la coja...Retener en un sitio sin tener salida...Cerrar documentos muy importantes en cerradura con llave en algún archivo...Encarcelar... Guardar algo bajo llave, trancar...Recogerse en clausura... Enjaular...Enclaustrar...
I punto: Poner presos en algún sitio... Ocultar algo sin darle sitio por donde salir...Cerrar, como una carta o una puerta... Guardar...No dejar salir... Tapar... Que yo encierre a uno de los nenes... Clausurar, apresar, tener preso... Encerrar a uno en la cárcel... Trancar... Cercar... Condenar las puertas.
0 puntos: Medio que tenemos para protegernos y nuestras casas... Mantener aislados...Algo que se guarda...Meter algo entre paredes...Mantener a uno en un sitio...Poner objeto dentro de otro, asegurar...ocultar... Esconder... Guardar silencio... Poner adentro.

14. DESAYUNO

2 puntos: Primer alimento que se toma en la mañana. Primera comida del día. Lo que se come por las mañanas...Una comida simple por la mañana... Lo que todos en general toman al levantarse...El pienso por la mañana.
I punto: Una merienda..Puede constar de café, pan, fruta, etc... Una parte de la nutrición del día...Aperitivo por la mañana.. Comer por las mañanas... Desayunarse por la mañana... Desayunar.

una persona aislada de los demás... Choza pequeña donde rezan o viven personas eclesiásticas... Capilla... Gruta.

1 punto: Santuario dedicado a algún santo... Sitio donde una persona hace retiro... Un refugio alejado de la urbanización... Iglesia, templo...

0 puntos: Especie de casa... Casucha... Una casa... Una cabaña... Sitio alto de algún campo donde hay tranquilidad... Valle... Edificio... Parte de un cementerio... Donde se guarda un objeto... Persona que vive aislada... Persona no civilizada.

30. OBSTRUIR

2 puntos: Estorbar el paso. Impedir el paso o acción. Interceptar. Interponerse... Poner obstáculos... Tapar el paso... Bloquear... Detener el libre paso... Cerrar el paso... Interferir... Tapar algún sitio o bocacalle... Interrumpir el tránsito.

1 punto: Ponerse en medio... No permitir... Obstruir la justicia... Interrumpir... Entorpecer... Oponerse... Hacer difícil el paso... No dejar hacer algo o no dejar terminar algo. Ejemplo: tapón en carretera...

0 puntos: Romper o destruir algo... Hacer algo, o ser causa de algo... Detener... Molestar... Desobedecer... Construir ... Intervenir.

31. INVENTAR

2 puntos: Hallar o descubrir con ingenio y estudio (una cosa nueva o no conocida). Crear. Imaginar (hechos falsos). Hacer algo nuevo por la cabeza de uno... Hacer algo que nadie ha hecho... Crear algo... Hacer algo original, de sus propias ideas... Hacer algo original, nuevo, propio...

1 punto: Descubrir algo... Hacer (producir) algo nuevo... Idear y producir.

0 puntos: Cuando uno hace cualquier experimento... Fabricar... Hacer algo... Producir, construir... Tratar de descubrir algo... Algo que se hace por la primera vez... Cuando inventa algo como una comida especial... Originar... Inventar un invento.

COMPOSICION DE OBJETOS

INSTRUCCIONES Presente siempre las figuras en el siguiente orden: el MANIQUI, el PERFIL, la MANO y el ELEFANTE.

Coloque las piezas de cada figura según se indica en los diagramas que aparecen en las páginas 69-70 ocultándoselas al sujeto. (La caja de materiales incluye una pantalla en forma de biombo para este propósito). Una vez que haya colocado las piezas del Maniquí, quite la pantalla, y diga: "Si estas piezas se juntan correctamente, formarán algo. Empiece, y júntelas lo más pronto que pueda".

Coloque las piezas de las otras figuras en sucesión, de acuerdo con los diagramas correspondientes, diciendo en cada caso: "Junte estas piezas lo más pronto que pueda". No permita que el sujeto vea los dibujos de las figuras ya terminadas que aparecen en las páginas siguientes:

LIMITE DE TIEMPO:	Maniquí y Perfil	120 segundos para cada figura
	Mano y Elefante	180 segundos para cada figura

Comience a tomar el tiempo tan pronto como termine de dar las instrucciones. Anote el tiempo exacto que el sujeto se tarde en juntar correctamente las partes de cada figura, ya que se acreditan puntuaciones adicionales por la rapidez con que el sujeto las complete dentro del límite de tiempo.

VALORACION Déle una puntuación de 2 por cada unión correcta que haga el sujeto dentro del límite de tiempo. (Pequeñas desviaciones, por ejemplo, separación en las uniones de no más de un cuarto de pulgada, o de medio centímetro, son aceptables). Las uniones que tienen que ser valoradas están señaladas por una "X" en los diagramas que aparecen en las páginas 69-70. Dé crédito por cada unión correcta aun cuando aparezca en un arreglo parcial. Por ejemplo, si el sujeto une las dos piezas que forman la oreja del Perfil pero sin colocarlas dentro de la cabeza, se da un punto; pero si se unen estas dos piezas correctamente a la cabeza, se dan 3 puntos. Igualmente, si el sujeto une correctamente dos piezas cualquiera de las que forman el Elefante, se le da un punto, aun cuando éstas no estén unidas con ninguna otra parte del cuerpo.

Los puntos que se dan por completar las figuras perfectamente, sin

TV... Llevar un mensaje... Hablar por teléfono... Hablar de un sitio a otro... Divulgar... Decir algo a otra persona.

0 puntos: Dejar conocer a otros algo... Oir por medio de cordones eléctricos... Decir... Decir lo mismo... Pasar por mediación del pensamiento... Un radio, un teléfono, un televisor.

20. INCOMPARABLE

2 puntos: Unico. Sin par.
Sin comparación... Que no se puede comparar... Que no tenga igual... inigualable... Que no tiene rival, lo mejor o lo más malo...

1 punto: Algo limitado o algo específico... Es cuando una persona no se puede comparar con otra... Que no se asemeja a alguien... Que no sea igual...El que pasó a todos. Que no se puede comparar con otro.

0 puntos: Es alguien que no le gusta cooperar...Cosa que es mejor que otra, o sea, que no son parecidas en nada... Que es casi parecido... Aquello que es mejor que otro... Es cuando se compara una cosa con otra... Dos personas que no se parecen... Inmejorable... Superior.

21. COMER

2 puntos: Tomar alimento. Tomar la comida principal del día.
Mascar y tragar el alimento.
Alimentarse... Ingerir alimentos...

1 punto: Corroer, desgastar, gastar, consumir... Comerse una ficha... Medio de sostenimiento... Sentir comezón física o moral... Ingerir... Tragar algo... Engullir.... Dirigir algo al estómago.

0 puntos: No se puede vivir sin comer... Lo que uno hace para mantenerse fuerte y estar saludable... Lo que uno va a comer, la comida... Algo que uno come.

22. ETERNO

2 puntos: De duración infinita. Lo que se halla fuera de la acción del tiempo - de lo que no tiene principio ni fin. Válido o existente en todos los tiempos. Eternal.

12. REGOCIJO

2 puntos: Júbilo...Alegría...Alborozo...Contento.
I punto : Gozo...Felicidad...Algo agradable...Sentir un gran placer por algo...Estar alegre o contento... Lo que sentimos cuando obtenemos éxito... Una persona que está alegre...Satisfacción...Uno se regocija si viene cualquier persona a la casa que hace tiempo uno no ha visto... Alegre, alegrarse...
0 puntos: Paz, tranquilidad...Sentirse orgulloso de algo... Es ser bueno...Halago... Algo que ha sido de nuestro agrado... Esmero...Entusiasmo... Se siente bien.

13. ENCERRAR

2 puntos: Meter a una persona o cosa en parte donde no puede salir, incluir (en el sentido figurado). Contener. Poner en un sitio para que nadie la coja...Retener en un sitio sin tener salida...Cerrar documentos muy importantes en cerradura con llave en algún archivo...Encarcelar... Guardar algo bajo llave, trancar...Recogerse en clausura... Enjaular...Enclaustrar...
I punto: Poner presos en algún sitio... Ocultar algo sin darle sitio por donde salir...Cerrar, como una carta o una puerta... Guardar...No dejar salir... Tapar... Que yo encierre a uno de los nenes... Clausurar, apresar, tener preso... Encerrar a uno en la cárcel... Trancar... Cercar... Condenar las puertas.
0 puntos: Medio que tenemos para protegernos y nuestras casas... Mantener aislados...Algo que se guarda...Meter algo entre paredes...Mantener a uno en un sitio...Poner objeto dentro de otro, asegurar...ocultar... Esconder... Guardar silencio... Poner adentro.

14. DESAYUNO

2 puntos: Primer alimento que se toma en la mañana. Primera comida del día. Lo que se come por las mañanas...Una comida simple por la mañana... Lo que todos en general toman al levantarse...El pienso por la mañana.
I punto: Una merienda..Puede constar de café, pan, fruta, etc... Una parte de la nutrición del día...Aperitivo por la mañana.. Comer por las mañanas... Desayunarse por la mañana... Desayunar.

25. REMONTAR

2 puntos: Ahuyentar o espantar. Elevar una cosa por el aire. Encumbrar. Echar nuevos pies o suelos (al calzado). Subir y volar muy alto (las aves). (Fig.) Subir hasta el origen de una cosa. Ascender... Subir... Alzar, levantar.

1 punto : Cuando está subiendo algo... Recordar tiempo remoto... Volver a montar... Alejarse... Emprender vuelo... Volver al principio de una cosa... Volver atrás, a lo viejo... Repechar... Trepar... Poner por alto... Elevar,(?)...

0 puntos: Volver a hacer algo... Viajar a algún lugar... Pasar de una etapa a otra... Vivir en el pasado... Pensar... Trasladarse... Llegar... Volver a empezar... Montar... Retornar... Montarse en alguna bicicleta, caballo, etc... Sobrepasar.

26. INDISPENSABLE

2 puntos: Que no se puede dispensar. Imprescindible. Algo sumamente necesario. Que no puede faltar... Algo que tiene que ser, que no puede dejarse... Que es un requisito... Que no se puede prescindir... Que no es dispensable.

1 punto: Que no se puede perdonar (de hacer algo que debe)... Cuando hacemos algo necesario ... Hace falta para obtener algo... Que no se puede sustituir... Que hace mucha falta... Algo necesario... Algo que tiene que hacer...

0 puntos: Que está seguro de algo... Provechoso, beneficioso... Algo único... Que no se puede aceptar... Inevitable... Algo que está bien hecho... Algo que no ha podido ser cumplido... Algo que no es necesario... Algo que se debe hacer... Que hace falta para algo... No hay disculpa... Sin remedio... Util...Que uno es dispensable.

27. REMORDIMIENTO

2 puntos: Pesar interno que queda después de ejecutar una mala acción. Pesar interno que surge como resultado de sentimiento de culpabilidad. Sentirse mal, ofuscado, después de haber hecho algo que no debía...Dolor en la conciencia por haber hecho algo malo... Sentimiento de culpa por algo que ha hecho mal... Pena que le cae a una persona por mala acción hecha.

1 punto: Cuando una persona se ha portado mal y está arrepentida de eso... Tener en la conciencia un remordimiento por haber hecho algo malo... Recuerdo de algo hecho que lo atormenta... Pensamiento que perdura por mucho tiempo si ha hecho algo malo... Nos molesta nuestra conciencia... Reproche que nos acusa de haber hecho algo malo...Algún pesar que uno siente... Sentimiento de conciencia... Arrepentimiento, pesar, culpabilidad... Sentirse acusado, perseguido... Cuando algo le está acusando de algo que ha hecho... Cuando uno está remordido... Sentimiento de culpabilidad.

0 puntos: Coraje ... Duda por algo que hizo mal... Sentimiento hacia alguien... Guardar rencor... Sentimiento de pena... Recuerdo de algún acto hecho mal... Mal humor... Problema...Algo grande que le hacen a uno y uno se queda con el sentimiento...Acusado...Rencor ...Resentimiento...Una persona que le ha dolido algo que le ha pasado... Arrepentimiento.

28. DESIGNAR

2 puntos: Formar designio de realizar (un trabajo). Denominar o determinar una (cosa o persona) por su nombre o rasgo distinto. Señalar (una persona o cosa) para determinado fin, Nombrar. Denominar.
Darle nombre a lugar o cosa... Elegir o dar un puesto u oficio a alguien... Asignar... Escoger a alguien o algo... Señalar...

1 punto : Mandar a hacer algo... Encomendar... Delegar... Dar algo para hacer... Elegir... Dar sitio... Determinar... Por ejemplo, designar fondos para una entidad u organización.

0 puntos: Resignación... Ordenar... Dar una orden... Mandar... No asignado... Distribuir algo... Darle algo a alguien... Persona o animal que tiene una tarea especial... Repartir... Entregar... No querer... Estar conforme.

29. ERMITA

2 puntos: Capilla o santuario situado generalmente en despoblado. Donde vive un ermitaño. .. Pequeña construcción donde se guardan cosas religiosas o se hacen liturgias... Donde vive

una persona aislada de los demás... Choza pequeña donde rezan o viven personas eclesiásticas... Capilla... Gruta.

1 punto: Santuario dedicado a algún santo... Sitio donde una persona hace retiro... Un refugio alejado de la urbanización... Iglesia, templo...

0 puntos: Especie de casa... Casucha... Una casa... Una cabaña... Sitio alto de algún campo donde hay tranquilidad... Valle... Edificio... Parte de un cementerio... Donde se guarda un objeto... Persona que vive aislada... Persona no civilizada.

30. OBSTRUIR

2 puntos: Estorbar el paso. Impedir el paso o acción. Interceptar. Interponerse... Poner obstáculos... Tapar el paso... Bloquear... Detener el libre paso... Cerrar el paso... Interferir... Tapar algún sitio o bocacalle... Interrumpir el tránsito.

1 punto: Ponerse en medio... No permitir... Obstruir la justicia... Interrumpir... Entorpecer... Oponerse... Hacer difícil el paso... No dejar hacer algo o no dejar terminar algo. Ejemplo: tapón en carretera...

0 puntos: Romper o destruir algo... Hacer algo, o ser causa de algo... Detener... Molestar... Desobedecer... Construir ... Intervenir.

31. INVENTAR

2 puntos: Hallar o descubrir con ingenio y estudio (una cosa nueva o no conocida). Crear. Imaginar (hechos falsos). Hacer algo nuevo por la cabeza de uno... Hacer algo que nadie ha hecho... Crear algo... Hacer algo original, de sus propias ideas... Hacer algo original, nuevo, propio...

1 punto: Descubrir algo... Hacer (producir) algo nuevo... Idear y producir.

0 puntos: Cuando uno hace cualquier experimento... Fabricar... Hacer algo... Producir, construir... Tratar de descubrir algo... Algo que se hace por la primera vez... Cuando inventa algo como una comida especial... Originar... Inventar un invento.

LISTA DE TABLAS

Tabla		Página
I	Muestra de Estandarización por Edad y Sexo	9
2	Muestra de Estandarización : Región por Zona de Residencia	10
3	Muestra de Estandarización por Edad, Sexo y Zona	11
4 A	Muestra de Estandarización por Sexo, Edad y Ocupación: Hombres	13
4 B	Muestra de Estandarización: por Sexo, Edad y Ocupación : Mujeres	14
5	Muestra de Estandarización por Edad, Sexo y Ocupación	15-17
6	Coeficientes de Confiabilidad y Errores Estándar de Medición (EEM) de las Subpruebas y Escalas Totales	20
7	Intercorrelaciones de las Subpruebas -Edades 16 - 19	22
8	Intercorrelaciones de l as Subpruebas Edades 25 - 34	23
9	Intercorrelaciones de la Subpruebas Edades 45-54	24
10	Promedios Aritméticos y Desviaciones Estándar de las Sumas de las Puntuaciones a Escala: Verbal, Manipulación y Total, por Grupos de Edad	26
11	Relación de los C I de la EIWA con Rangos Percentiles	27
12	Clasificaciones de Inteligencia	27

5. Asegúrese de que la puntuación bruta de cada subprueba se transfiera al espacio correspondiente en la Sección de resumen de los resultados que aparece en la portada de la Hoja de anotaciones.

6. Evite errores al convertir las puntuaciones brutas a las puntuaciones a escala y también al convertir éstas en C I. Repita los pasos envueltos en este proceso para corregir posibles errores en la lectura de las tablas.

7. Verifique la edad del sujeto comparando la edad informada por el sujeto con la diferencia entre la fecha de administración de la prueba y la fecha de nacimiento, o consultando récords confiables.

8. Revise toda la información transferida, las computaciones y la conversión de puntuaciones para más exactitud.

Conversión de las puntuaciones brutas a puntuaciones a escala

La puntuación bruta en cada subprueba es la suma del número de puntos obtenidos en cada ítem. Estas puntuaciones brutas se transfieren a la sección de resumen de los resultados que se provee en la portada de la Hoja de anotaciones. Hacia la izquierda de ésta hay una Tabla de equivalencias a puntuaciones a escala (que aparece también en la página 103 de este Manual) que se usa para convertir las puntuaciones brutas en puntuaciones a escala de todos los sujetos, independientemente de su edad y sexo. La puntuación bruta que un sujeto obtiene en una subprueba se localiza en la columna de la tabla correspondiente a esa subprueba. El examinador podrá localizar la puntuación a escala correspondiente a cierta puntuación bruta leyendo horizontalmente hacia la izquierda o hacia la derecha de la columna en la Tabla. Esta puntuación a escala se anota en el espacio correspondiente en la Sección de resumen de los resultados a la derecha de la puntuación bruta. Al terminar este procedimiento con todas las subpruebas, la Sección de resumen de los resultados muestra una columna de puntuaciones brutas y al lado de ésta, una de puntuaciones a escala. De aquí en adelante no es necesario usar las puntuaciones brutas, pues las puntuaciones a escala son más significativas y permiten tratamiento aritmético.[1]

---

[1] Se provee a la derecha de esta columna un espacio adicional que el examinador puede usar para colocar las desviaciones de puntuación del sujeto en cada subprueba o cualquier otra anotación que se quiera hacer.

Para cargar...De embarcar ... En barco se puede viajar ...Algo de madera que va por el agua... Para llevar pasajeros, carga...Caminar por el mar ... Canoa, lancha, vapor, bote...Un vehículo.

0 puntos: Hecho de metal ...Medio más antiguo que usó el hombre...Es muy cómodo... Lo que los países usan para comercio o para la defensa nacional ... Embarque

6. ESCUELA

2 puntos: Establecimiento público o privado donde se dá cualquier género de instrucción y especialmente la primaria... . Plantel de enseñanza...Centro de enseñanza ... Institución donde se enseña...Aula.

1 punto : Donde educan a los niños...Que uno estudia en ella... Donde uno aprende...Donde los niños van a estudiar... Sitio donde se estudia...Le dan clases a uno...Colegio.

0 puntos: Estudio...Enseñanza...Educación...Un salón.

7. MESA

2 puntos: Mueble compuesto por una tabla lisa sostenida por uno o varios pies..
Un mueble... Tabla lisa con cuatro patas.

1 punto: Mesa en que se come...Para sentarse a escribir... Donde uno come.

0 puntos : Un objeto de la casa...Quitar o poner la mesa... Lo que se usa para comer ...Sentarse a la mesa.

8. OCULTAR

2 puntos: Esconder. Poner en un lugar o sitio retirado o secreto para que no sea encontrado fácilmente. Impedir que sea vista. Encubrir. Poner fuera de la vista. No permitir que otros vean...Guardar algo en secreto, tapar ... Reservar...Disimular (sentimiento o pensamiento)

1 punto : Mentira, no decir la verdad...Tener oculto...Secreto... Algo que uno vea y lo oculta, no diría nada...Negar una cosa...No divulgar...Escondido...Cuando uno hace algo malo y no quiere que el papá lo sepa.

Algo que está mal puesto... Cuando uno requiere algo ... Es seguir algo... Grupo de personas que sirven a un alto porcentaje. Grupo de personas que van a participar en alguna actividad... Celebrar una boda... Que no está enojado... Trión... Tribu... Boda, entierro.

EIWA

ESCALA DE INTELIGENCIA
WECHSLER PARA ADULTOS

Nombre ______________________________

Fecha de Nacimiento ________ Día ________ Mes ________ Año Edad ________ Sexo ________ Estado Civil ________

Nacionalidad ______________________ Color ________ Examinador ________

Sitio de Examen ______________________ Fecha ________

Ocupación ______________________ Educación ______________________

TABLA DE EQUIVALENCIAS EN PUNTUACIONES A ESCALA

PUNTUACION A ESCALA	PUNTUACION BRUTA											
	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario	Dígito-Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos	Puntuación a Escala
19	32	27-30	--	28-30	16-17	73-80	93-110	25	--	47-48	43-44	19
18	31	25-26	16	26-27	14-15	70-72	78-92	24	46	45-46	41-42	18
17	30	24	--	24-25	13	66-69	72-78	--	45	43-44	39-40	17
16	29	22-23	15	22-23	12	61-65	67-71	23	43-44	41-42	36-38	16
15	26-28	20-21	14	20-21	--	57-60	63-66	22	39-42	39-40	33-35	15
14	23-25	18-19	12-13	18-19	11	49-56	60-62	20-21	35-38	37-38	30-32	14
13	20-22	16-17	11	14-17	10	43-48	54-59	19	33-34	34-36	27-29	13
12	17-19	14-15	9-10	11-13	9	37-42	47-53	17-18	24-32	30-33	23-26	12
11	14-16	12-13	8	9-10	--	30-36	41-46	15-16	24-28	25-29	20-22	11
10	11-13	10-11	7	7-8	8	24-29	35-40	13-14	19-23	18-24	15-19	10
9	9-10	8-9	6	6	7	19-23	29-34	11-12	13-18	13-17	12-14	9
8	8	6-7	--	5	--	16-18	22-28	9-10	10-12	9-12	9-11	8
7	7	5	5	4	6	12-15	14-21	7-8	7-9	7-8	6-8	7
6	6	4	4	3	5	9-11	10-13	5-6	5-6	5-6	4-5	6
5	5	--	--	2	4	5-8	7-9	4	4	3-4	3	5
4	4	3	3	1	3	3-4	4-6	3	3	2	2	4
3	3	--	--	--	--	2	1-3	2	2	1	--	3
2	2	2	2	0	--	1	0	1	1	--	1	2
1	1	1	1	--	2	--	--	--	0	0	--	1
0	0	0	0	--	0	0	--	0	--	--	0	0

RESUMEN

Subprueba	Puntuación Bruta	Puntuación a Escala	
Información			
Comprensión			
Aritmética			
Analogías			
Repetición de dígitos			
Vocabulario			
Puntuación Verbal			
Dígito-Símbolo			
Dibujos para completar			
Diseños con cubos			
Ordenamiento de dibujos			
Composición de objetos			
Puntuación-Manipulación			
Puntuación Total			

Puntuación Verbal ________ C I ________

Puntuación de Manipulación ________ C I ________

Puntuación Total ________ C I ________

	Puntuación 1 - 0	1. INFORMACION
1. alas		
2. pelota		
3. yerba		
4. azúcar		
5. meses		
6. termómetro		
7. idiomas		
8. caminar		
9. sangre		
10. metal		
11. goma		
12. Rubén Darío		
13. corazón		
14. Isabel La Católica		
15. objeto		
16. Vaticano		
17. Italia		
18. Río de Janeiro		
19. fósforo		
20. café		
21. Quijote		
22. semana		
23. países		
24. América		
25. relámpago		
26. levadura		
27. ropa		
28. vasos sanguíneos		
29. civilizaciones		
30. Ilíada		
31. Corán		
32. Egipto		

2. COMPRENSION		Puntuación 2, 1-0
1. ropa		
2. alimentos		
3. sobre		
4. ancianos		
5. cine		
6. niño		
7. terreno		
8. brasas		
9. impuestos		
10. perro		
11. gobierno		
12. igualdad		
13. sordos		
14. bosque		
15. golondrina		

3. ARITMETICA			
	B σ M	Tiempo	Puntuación 1-0
1. 15"			
2. 15"			
3. 15"			
4. 15"			
5. 15"			
6. 15"			
7. 30"			
8. 30"			
9. 30"			
10. 30"			
11. 30"			
12. 30"			
13. 60"			
14. 60"			
15. 60"			
16. 60"			

4. ANALOGIAS		Puntuación 2-1-0
1. lápiz-pluma		
2. revólver -escopeta		
3. cordón - soga		
4. baúl - maleta		
5. hacha - serrucho		
6. perro - toro		
7. oro - plata		
8. avión - caballo		
9. pantalón - zapato		
10. aire - agua		
11. ojo - oído		
12. norte -oeste		
13. sangre -savia		
14. poesía - estatua		
15. día - litro		

5. REPETICION DE DIGITOS	Puntuación
Repetición Directa	Circule
5-8-2	3
6-9-4	3
6-4-3-9	4
7-2-8-6	4
4-2-7-3-1	5
7-5-8-3-6	5
6-1-9-4-7-3	6
3-9-2-4-8-7	6
5-9-1-7-4-2-8	7
4-1-7-9-3-8-6	7
5-8-1-9-2-6-4-7	8
3-8-2-9-5-1-7-4	8
2-7-5-8-6-2-5-8-4	9
7-1-3-9-4-2-5-6-8	9
Repetición Inversa	Circule
2-4	2
5-8	2
6-2-9	3
4-1-5	3
3-2-7-9	4
4-9-6-8	4
1-5-2-8-6	5
6-1-8-4-3	5
5-3-9-4-1-8	6
7-2-4-8-5-6	6
8-1-2-9-3-6-5	7
4-7-3-9-1-2-8	7
9-4-3-7-6-2-5-8	8
7-2-8-1-9-6-5-3	8

D______ / I______ = ______

número mas alto circulado

	Puntuación 2,1-0	6. VOCABULARIO
1. finalizar		
2. gato		
3. apresurar		
4. comenzar		
5. barco		
6. escuela		
7. mesa		
8. ocultar		
9. invierno		
10. sombrero		
11. reunir		
12. regocijo		
13. encerrar		
14. desayuno		
15. doméstico		
16. admitir		
17. meditar		
18. caverna		
19. transmitir		
20. incomparable		

	Puntuación 2,1-0	VOCABULARIO (CONT.)
21. comer		
22. eterno		
23. firmeza		
24. consumir		
25. remontar		
26. indispensable		
27. remordimiento		
28. designar		
29. ermita		
30. obstruir		
31. inventar		
32. abolir		
33. palpar		
34. regularizar		
35. fugaz		
36. espléndido		
37. osado		
38. insensato		
39. rebasar		
40. séquito		

7. DIGITOS
SIMBOLOS

1	2	3	4	5	6
⊥	—	○	||	✕	∧

PUNTUACIÓN

EJERCICIOS DE PRÁCTICA

2	1	3	1	2	4	6	1	5	4	2	1	3	2	1	4	2	3	5	2	3	1	4	6	3
1	5	4	2	4	6	3	5	3	2	4	5	4	6	3	2	1	3	2	6	5	3	4	6	3
6	2	5	1	5	2	6	3	5	4	6	5	4	1	4	3	5	2	6	1	5	4	6	3	4
5	2	6	1	3	6	4	6	3	5	1	5	2	4	5	2	1	4	3	6	3	1	4	1	6
6	2	5	4	1	6	3	6	2	1	5	6	4	3	1	5	4	3	5	4	6	3	1	4	2

8. DIBUJOS para COMPLETAR	Puntuación 1 - 0	Tiempo
1. ala		
2. ojo		
3. pantalla		
4. rabo		
5. zapato		
6. abertura para llave		
7. dientes		
8. rueda		
9. palillo		
10. fuego		
11. clavija		
12. nariz		
13. lengueta		
14. media		
15. puente		
16. portón		
17. cadena		
18. imagen del brazo		
19. dedo		
20. el uno del diez		
21. horquilla		
22. pluma		
23. sombra		
24. chimenea		
25. ceja		

9. DISEÑOS CON CUBOS

		Tiempo	1-0	PUNTUACION			
A.	60"	1 2					
1.	60"	1 2		0	2	4	
2.	60"	1 2		0	2		
3.	60"			0	4		
4.	60"			0	4		
5.	60"			0	4		
6.	60"			0	4		
7.	60"			0	4		
8.	120"			0	4	46-70 5	1-45 6
9.	120"			0	4	61-80 5	1-60 6
10.	120"			0	4	61-80 5	1-60 6

10. ORDENAMIENTO DE DIBUJOS

		Tiempo	Orden	PUNTUACION				
1. ave	60"	1 2		0	2	4	wxy	
2. casa	60"	1 2		0	2	4	pat	
3. jira	60"			0		4	efgh	
4. caída	60"			0		4	sube	
5. asalto	60"			0		4	abcd	
6. otra	60"			0		4	otra	
7. serenata	120"			0 canto		26-120 4	16-25 5	1-15 6
8. cazador	120"			0 fusila		26-120 4	16-25 5	1-15 6
9. pibe	120"			0 lucha		26-120 4	16-25 5	1-15 6
10. galanteo	120"			0 Leisa eilsa	2	26-120 4 Elisa	16-25 5	1-15 6

11. COMPOSICION DE OBJETOS

		Tiempo	Puntuación								
1. maniquí	120"		0 1 2 3 4	5	16-20 6	11-15 7	1-10 8				
2. perfil	120"		0 1 2 3 4	5	6	7	8	9	36-45 11	26-35 12	1-25 13
3. mano	180"		0 1 2 3 4	5	6	7	41-50 9	31-40 10	1-30 11		
4. elefante	180"		0 1 2 3 4	5	6	7	8	31-50 10	21-30 11	1-20 12	

PUNTUACIONES BRUTAS Y SUS EQUIVALENTES EN PUNTUACIONES A ESCALA

SUMAS DE LAS PUNTUACIONES A ESCALA Y SUS EQUIVALENTES EN COCIENTES DE INTELIGENCIA

Tabla 15. Puntuaciones Brutas y sus Equivalentes en Puntuaciones a Escala

	VERBAL							MANIPULACION				
	Información	Comprensión	Aritmética	Analogías	Repetición de dígitos	Vocabulario		Dígito Símbolo	Dibujos para completar	Diseños con cubos	Ordenamiento de dibujos	Composición de objetos
Promedio	10.03	9.93	10.08	9.97	10.03	9.95		9.93	9.93	9.88	9.87	9.96
Desviación Estándar	3.05	2.93	3.07	3.22	2.96	3.13		3.16	3.02	3.14	3.10	3.00
Puntuación a Escala							Puntuación a Escala					
19	32	27-30	--	28-30	16-17	73-80	19	93-110	25	--	47-48	43-44
18	31	25-26	16	26-27	14-15	70-72	18	79-92	24	46	45-46	41-42
17	30	24	--	24-25	13	66-69	17	72-78	--	45	43-44	39-40
16	29	22-23	15	22-23	12	61-65	16	67-71	23	43-44	41-42	36-38
15	26-28	20-21	14	20-21	--	57-60	15	63-66	22	39-42	39-40	33-35
14	23-25	18-19	12-13	18-19	11	49-56	14	60-62	20-21	35-38	37-38	30-32
13	20-22	16-17	11	14-17	10	43-48	13	54-59	19	33-34	34-36	27-29
12	17-19	14-15	9-10	11-13	9	37-42	12	47-53	17-18	29-32	30-33	23-26
11	14-16	12-13	8	9-10	--	30-36	11	41-46	15-16	24-28	25-29	20-22
10	11-13	10-11	7	7-8	8	24-29	10	35-40	13-14	19-23	18-24	15-19
9	9-10	8-9	6	6	7	19-23	9	29-34	11-12	13-18	13-17	12-14
8	8	6-7	--	5	--	16-18	8	22-28	9-10	10-12	9-12	9-11
7	7	5	5	4	6	12-15	7	14-21	7-8	7-9	7-8	6-8
6	6	4	4	3	5	9-11	6	10-13	5-6	5-6	5-6	4-5
5	5	--	--	2	4	5-8	5	7-9	4	4	3-4	3
4	4	3	3	1	3	3-4	4	4-6	3	3	2	2
3	3	--	--	--	--	2	3	1-3	2	2	1	--
2	2	2	2	0	--	1	2	0	1	1	--	1
1	1	1	1	--	2	--	1	--	--	0	0	--
0	0	0	0	--	0	0	0	--	0	--	--	0

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 16 - 19

VERBAL

Puntuación	C I	Puntuación	C I
113	152	57	98
112	151	56	97
111	150	55	96
110	149	54	95
109	148	53	94
108	148	52	93
107	147	51	92
106	146	50	91
105	145	49	90
104	144	48	89
103	143	47	89
102	142	46	88
101	141	45	87
100	140	44	86
99	139	43	85
98	138	42	84
97	137	41	83
96	136	40	82
95	135	39	81
94	134	38	80
93	133	37	79
92	132	36	78
91	131	35	77
90	130	34	76
89	129	33	75
88	128	32	74
87	127	31	73
86	126	30	72
85	125	29	71
84	124	28	70
83	123	27	69
82	122	26	68
81	121	25	67
80	120	24	66
79	119	23	65
78	118	22	64
77	118	21	63
76	117	20	62
75	116	19	61
74	115	18	60
73	114	17	59
72	113	16	59
71	112	15	58
70	111	14	57
69	110	13	56
68	109	12	55
67	108	11	54
66	107	10	53
65	106	9	52
64	105	8	51
63	104	7	50
62	103	6	49
61	102	5	48
60	101	4	47
59	100	3	46
58	99	2	45

MANIPULACION

Puntuación	C I	Puntuación	CI
94	147	48	95
93	145	47	93
92	144	46	92
91	143	45	91
90	142	44	90
89	141	43	89
88	140	42	88
87	139	41	87
86	138	40	86
85	136	39	84
84	135	38	83
83	134	37	82
82	133	36	81
81	132	35	80
80	131	34	79
79	130	33	78
78	129	32	77
77	127	31	75
76	126	30	74
75	125	29	73
74	124	28	72
73	123	27	71
72	122	26	70
71	121	25	69
70	119	24	68
69	118	23	66
68	117	22	65
67	116	21	64
66	115	20	63
65	114	19	62
64	113	18	61
63	112	17	60
62	110	16	58
61	109	15	57
60	108	14	56
59	107	13	55
58	106	12	54
57	105	11	53
56	104	10	52
55	103	9	51
54	101	8	49
53	100	7	48
52	99	6	47
51	98	5	46
50	97	4	45
49	96		

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 16 - 19

ESCALA TOTAL

Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
		158	125	107	97	56	70
		157	125	106	97	55	69
207	152	156	124	105	96	54	69
206	151	155	124	104	96	53	68
205	151	154	123	103	95	52	68
204	150	153	123	102	95	51	67
203	150	152	122	101	94	50	66
202	149	151	121	100	94	49	66
201	149	150	121	99	93	48	65
200	148	149	120	98	93	47	65
199	148	148	120	97	92	46	64
198	147	147	119	96	91	45	64
197	146	146	119	95	91	44	63
196	146	145	118	94	90	43	63
195	145	144	118	93	90	42	62
194	145	143	117	92	89	41	62
193	144	142	117	91	89	40	61
192	144	141	116	90	88	39	60
191	143	140	115	89	88	38	60
190	143	139	115	88	87	37	59
189	142	138	114	87	87	36	59
188	142	137	114	86	86	35	58
187	141	136	113	85	86	34	58
186	140	135	113	84	85	33	57
185	140	134	112	83	84	32	57
184	139	133	112	82	84	31	56
183	139	132	111	81	83	30	56
182	138	131	111	80	83	29	55
181	138	130	110	79	82	28	54
180	137	129	109	78	82	27	54
179	137	128	109	77	81	26	53
178	136	127	108	76	81	25	53
177	136	126	108	75	80	24	52
176	135	125	107	74	80	23	52
175	134	124	107	73	79	22	51
174	134	123	106	72	78	21	51
173	133	122	106	71	78	20	50
172	133	121	105	70	77	19	50
171	132	120	105	69	77	18	49
170	132	119	104	68	76	17	49
169	131	118	103	67	76	16	48
168	131	117	103	66	75	15	47
167	130	116	102	65	75	14	47
166	130	115	102	64	74	13	46
165	129	114	101	63	73	12	46
164	128	113	101	62	73	11	45
163	128	112	100	61	72	10	45
162	127	111	100	60	72	9	44
161	127	110	99	59	71	8	44
160	126	109	99	58	71	7	43
159	126	108	98	57	70	6	43

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 20 - 24

VERBAL

Puntuación	C I	Puntuación	C I
113	152	57	98
112	151	56	97
111	150	55	96
110	149	54	95
109	148	53	94
108	148	52	93
107	147	51	92
106	146	50	91
105	145	49	90
104	144	48	89
103	143	47	89
102	142	46	88
101	141	45	87
100	140	44	86
99	139	43	85
98	138	42	84
97	137	41	83
96	136	40	82
95	135	39	81
94	134	38	80
93	133	37	79
92	132	36	78
91	131	35	77
90	130	34	76
89	129	33	75
88	128	32	74
87	127	31	73
86	126	30	72
85	125	29	71
84	124	28	70
83	123	27	69
82	122	26	68
81	121	25	67
80	120	24	66
79	119	23	65
78	118	22	64
77	118	21	63
76	117	20	62
75	116	19	61
74	115	18	60
73	114	17	59
72	113	16	59
71	112	15	58
70	111	14	57
69	110	13	56
68	109	12	55
67	108	11	54
66	107	10	53
65	106	9	52
64	105	8	51
63	104	7	50
62	103	6	49
61	102	5	48
60	101	4	47
59	100	3	46
58	99	2	45

MANIPULACION

Puntuación	C I	Puntuación	C I
94	149	48	97
93	148	47	96
92	147	46	95
91	145	45	93
90	144	44	92
89	143	43	91
88	142	42	90
87	141	41	89
86	140	40	88
85	139	39	87
84	137	38	85
83	136	37	84
82	135	36	83
81	134	35	82
80	133	34	81
79	132	33	80
78	131	32	79
77	130	31	78
76	128	30	76
75	127	29	75
74	126	28	74
73	125	27	73
72	124	26	72
71	123	25	71
70	122	24	70
69	121	23	69
68	119	22	67
67	118	21	66
66	117	20	65
65	116	19	64
64	115	18	63
63	114	17	62
62	113	16	61
61	111	15	60
60	110	14	58
59	109	13	57
58	108	12	56
57	107	11	55
56	106	10	54
55	105	9	53
54	104	8	52
53	102	7	50
52	101	6	49
51	100	5	48
50	99	4	47
49	98		

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 20-24

ESCALA TOTAL

Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
		158	126	107	99	56	71
		157	126	106	98	55	70
207	153	156	125	195	97	54	70
206	152	155	125	194	97	53	69
205	152	154	124	193	96	52	69
204	151	153	124	192	96	51	68
203	151	152	123	191	95	50	64
202	150	151	122	100	95	49	64
201	150	150	122	99	94	48	66
200	149	149	121	98	94	47	66
199	149	148	121	97	93	46	65
198	148	147	120	96	93	45	65
197	147	146	120	95	92	44	64
196	147	145	119	94	91	43	64
195	146	144	119	93	91	42	63
194	146	143	118	92	90	41	63
193	145	142	118	91	90	40	62
192	145	141	117	90	89	39	62
191	144	140	116	89	89	38	61
190	144	139	116	88	88	37	60
189	143	138	115	87	88	36	60
188	143	137	115	86	87	35	59
187	142	136	114	85	87	34	59
186	142	135	114	84	86	33	58
185	141	134	113	83	85	32	58
184	140	133	113	82	85	31	57
183	140	132	112	81	84	30	57
182	139	131	112	80	84	29	56
181	139	130	111	79	83	28	56
180	138	129	110	78	83	27	55
179	138	128	110	77	82	26	54
178	137	127	109	76	82	25	54
177	137	126	109	75	81	24	53
176	136	125	108	74	81	23	53
175	136	124	108	73	80	22	52
174	135	123	107	72	79	21	52
173	134	122	107	71	79	20	51
172	134	121	106	70	78	19	51
171	133	120	106	69	78	18	50
170	133	119	105	68	77	17	50
169	132	118	105	67	77	16	49
168	132	117	104	66	76	15	48
167	131	116	193	65	76	14	48
166	131	115	193	64	75	13	47
165	130	114	192	63	75	12	47
164	130	113	192	62	74	11	46
163	129	112	101	61	73	10	46
162	128	111	101	60	73	9	45
161	128	100	100	59	72	8	45
160	127	109	100	58	72	7	44
159	127	108	99	57	71	6	44

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 25 - 34

VERBAL

Puntuación	C I	Puntuación	CI
113	151	57	96
112	150	56	95
111	149	55	94
110	148	54	93
109	147	53	93
108	146	52	92
107	145	51	91
106	144	50	90
105	143	49	89
104	142	48	88
103	141	47	87
102	140	46	86
101	139	45	85
100	138	44	84
99	137	43	83
98	136	42	82
97	135	41	81
96	134	40	80
95	133	39	79
94	132	38	78
93	131	37	77
92	130	36	76
91	129	35	75
90	128	34	74
89	127	33	73
88	126	32	72
87	125	31	71
86	124	30	70
85	123	29	69
84	122	28	68
83	122	27	67
82	121	26	66
81	120	25	65
80	119	24	64
79	118	23	63
78	117	22	63
77	116	21	62
76	115	20	61
75	114	19	60
74	113	18	59
73	112	17	58
72	111	16	57
71	110	15	56
70	109	14	55
69	108	13	54
68	107	12	53
67	106	11	52
66	105	10	51
65	104	9	50
64	103	8	49
63	102	7	48
62	101	6	47
61	100	5	46
60	99	4	45
59	98	3	44
58	97	2	43

MANIPULACION

Puntuación	CI	Puntuación	CI
94	151	48	99
93	150	47	98
92	149	46	97
91	148	45	96
90	147	44	95
89	146	43	94
88	144	42	92
87	143	41	91
86	142	40	90
85	141	39	89
84	140	38	88
83	139	37	87
82	138	36	86
81	136	35	84
80	135	34	83
79	134	33	82
78	133	32	81
77	132	31	80
76	131	30	79
75	130	29	78
74	129	28	77
73	127	27	75
72	126	26	74
71	125	25	73
70	124	24	72
69	123	23	71
68	122	22	70
67	121	21	69
66	120	20	68
65	118	19	66
64	117	18	65
63	116	17	64
62	115	16	63
61	114	15	62
60	113	14	61
59	112	13	60
58	110	12	59
57	109	11	57
56	108	10	56
55	107	9	55
54	106	8	54
53	105	7	53
52	104	6	52
51	103	5	51
50	101	4	49
49	100		

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 25 - 34

ESCALA TOTAL

Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
		158	126	107	99	56	71
		157	126	106	98	55	70
207	153	156	125	105	98	54	70
206	153	155	125	104	97	53	69
205	152	154	124	103	96	52	69
204	151	153	124	102	96	51	68
203	151	152	123	101	95	50	68
202	150	151	123	100	95	49	67
201	150	150	122	99	94	48	67
200	149	149	122	98	94	47	66
199	149	148	121	97	93	46	65
198	148	147	120	96	93	45	65
197	148	146	120	95	92	44	64
196	147	145	119	94	92	43	64
195	147	144	119	93	91	42	63
194	146	143	118	92	90	41	63
193	145	142	118	91	90	40	62
192	145	141	117	90	89	39	62
191	144	140	117	89	89	38	61
190	144	139	116	88	88	37	61
189	143	138	116	87	88	36	60
188	143	137	115	86	87	35	59
187	142	136	114	85	87	34	59
186	142	135	114	84	86	33	58
185	141	134	113	83	86	32	58
184	141	133	113	82	85	31	57
183	140	132	112	81	85	30	57
182	139	131	112	80	84	29	56
181	139	130	111	79	83	28	56
180	138	129	111	78	83	27	55
179	138	128	110	77	82	26	55
178	137	127	110	76	82	25	54
177	137	126	109	75	81	24	53
176	136	125	108	74	81	23	53
175	136	124	108	73	80	22	52
174	135	123	107	72	80	21	52
173	135	122	107	71	79	20	51
172	134	121	106	70	79	19	51
171	133	120	106	69	78	18	50
170	133	119	105	68	77	17	50
169	132	118	105	67	77	16	49
168	132	117	104	66	76	15	49
167	131	116	104	65	76	14	48
166	131	115	103	64	75	13	48
165	130	114	102	63	75	12	47
164	130	113	102	62	74	11	46
163	129	112	101	61	74	10	46
162	129	111	101	60	73	9	45
161	128	110	100	59	73	8	45
160	128	109	100	58	72	7	44
159	127	108	99	57	71	6	44

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 35 - 44

VERBAL				MANIPULACION			
Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
113	149	57	95	94	154	48	102
112	148	56	94	93	153	47	101
111	147	55	93	92	152	46	100
110	146	54	92	91	151	45	99
109	145	53	91	90	150	44	98
108	144	52	90	89	148	43	97
107	143	51	89	88	147	42	95
106	142	50	88	87	146	41	94
105	141	49	87	86	145	40	93
104	140	48	86	85	144	39	92
103	139	47	85	84	143	38	91
102	138	46	84	83	142	37	90
101	137	45	83	82	141	36	89
100	136	44	82	81	139	35	87
99	136	43	81	80	138	34	86
98	135	42	80	79	137	33	85
97	134	41	79	78	136	32	84
96	133	40	78	77	135	31	83
95	132	39	77	76	134	30	82
94	131	38	77	75	133	29	81
93	130	37	76	74	132	28	80
92	129	36	75	73	130	27	78
91	128	35	74	72	129	26	77
90	127	34	73	71	128	25	76
89	126	33	72	70	127	24	75
88	125	32	71	69	126	23	74
87	124	31	70	68	125	22	73
86	123	30	69	67	124	21	72
85	122	29	68	66	122	20	71
84	121	28	67	65	121	19	69
83	120	27	66	64	120	18	68
82	119	26	65	63	119	17	67
81	118	25	64	62	118	16	66
80	117	24	63	61	117	15	65
79	116	23	62	60	116	14	64
78	115	22	61	59	115	13	63
77	114	21	60	58	113	12	61
76	113	20	59	57	112	11	60
75	112	19	58	56	111	10	59
74	111	18	57	55	110	9	58
73	110	17	56	54	109	8	57
72	109	16	55	53	108	7	56
71	108	15	54	52	107	6	55
70	107	14	53	51	106	5	54
69	107	13	52	50	104	4	52
68	106	12	51	49	103		
67	105	11	50				
66	104	10	49				
65	103	9	48				
64	102	8	48				
63	101	7	47				
62	100	6	46				
61	99	5	45				
60	98	4	44				
59	97	3	43				
58	96	2	42				

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 35 - 44

ESCALA TOTAL

Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
		158	127	107	99	56	71
		157	126	106	99	55	71
207	154	156	126	105	98	54	70
206	153	155	125	104	98	53	70
205	153	154	125	103	97	52	69
204	152	153	124	102	97	51	69
203	151	152	124	101	96	50	68
202	151	151	123	100	95	49	68
201	150	150	123	99	95	48	67
200	150	149	122	98	94	47	67
199	149	148	122	97	94	46	66
198	149	147	121	96	93	45	66
197	148	146	120	95	93	44	65
196	148	145	120	94	92	43	64
195	147	144	119	93	92	42	64
194	147	143	119	92	91	41	63
193	146	142	118	91	91	40	63
192	145	141	118	90	90	39	62
191	145	140	117	89	89	38	62
190	144	139	117	88	89	37	61
189	144	138	116	87	88	36	61
188	143	137	116	86	88	35	60
187	143	136	115	85	87	34	60
186	142	135	114	84	87	33	59
185	142	134	114	83	86	32	58
184	141	133	113	82	86	31	58
183	141	132	113	81	85	30	57
182	140	131	112	80	85	29	57
181	140	130	112	79	84	28	56
180	139	129	111	78	83	27	56
179	138	128	111	77	83	26	55
178	138	127	110	76	82	25	55
177	137	126	110	75	82	24	54
176	137	125	109	74	81	23	54
175	136	124	109	73	81	22	53
174	136	123	108	72	80	21	52
173	135	122	107	71	80	20	52
172	135	121	107	70	79	19	51
171	134	120	106	69	79	18	51
170	134	119	106	68	78	17	50
169	133	118	105	67	77	16	50
168	133	117	105	66	77	15	49
167	132	116	104	65	76	14	49
166	131	115	104	64	76	13	48
165	131	114	103	63	75	12	48
164	130	113	103	62	75	11	47
163	130	112	102	61	74	10	46
162	129	111	101	60	74	9	46
161	129	110	101	59	73	8	45
160	128	109	100	58	73	7	45
159	128	108	100	57	72	6	44

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 45 - 54

VERBAL				MANIPULACION			
Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
113	150	57	96	94	160	48	108
112	149	56	95	93	159	47	107
111	148	55	94	92	158	46	106
110	147	54	93	91	157	45	105
109	147	53	92	90	156	44	104
108	146	52	91	89	155	43	103
107	145	51	90	88	153	42	101
106	144	50	89	87	152	41	100
105	143	49	88	86	151	40	99
104	142	48	88	85	150	39	98
103	141	47	87	84	149	38	97
102	140	46	86	83	148	37	96
101	139	45	85	82	147	36	95
100	138	44	84	81	145	35	94
99	137	43	83	80	144	34	92
98	136	42	82	79	143	33	91
97	135	41	81	78	142	32	90
96	134	40	80	77	141	31	89
95	133	39	79	76	140	30	88
94	132	38	78	75	139	29	87
93	131	37	77	74	138	28	86
92	130	36	76	73	136	27	84
91	129	35	75	72	135	26	83
90	128	34	74	71	134	25	82
89	127	33	73	70	133	24	81
88	126	32	72	69	132	23	80
87	125	31	71	68	131	22	79
86	124	30	70	67	130	21	78
85	123	29	69	66	129	20	77
84	122	28	68	65	127	19	75
83	121	27	67	64	126	18	74
82	120	26	66	63	125	17	73
81	119	25	65	62	124	16	72
80	118	24	64	61	123	15	71
79	118	23	63	60	122	14	70
78	117	22	62	59	121	13	69
77	116	21	61	58	120	12	68
76	115	20	60	57	118	11	66
75	114	19	59	56	117	10	65
74	113	18	59	55	116	9	64
73	112	17	58	54	115	8	63
72	111	16	57	53	114	7	62
71	110	15	56	52	113	6	61
70	109	14	55	51	112	5	60
69	108	13	54	50	110	4	59
68	107	12	53	49	109		
67	106	11	52				
66	105	10	51				
65	104	9	50				
64	103	8	49				
63	102	7	48				
62	101	6	47				
61	100	5	46				
60	99	4	45				
59	98	3	44				
58	97	2	43				

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 45 - 54

ESCALA TOTAL

Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
		158	131	107	103	56	75
		157	130	106	102	55	75
207	157	156	130	105	102	54	74
206	157	155	129	104	101	53	74
205	156	154	128	103	101	52	73
204	156	153	128	102	100	51	72
203	155	152	127	101	100	50	72
202	155	151	127	100	99	49	71
201	154	150	126	99	99	48	71
200	154	149	126	98	98	47	70
199	153	148	125	97	97	46	70
198	152	147	125	96	97	45	69
197	152	146	124	95	96	44	69
196	151	145	124	94	96	43	68
195	151	144	123	93	95	42	68
194	150	143	123	92	95	41	67
193	150	142	122	91	94	40	66
192	149	141	121	90	94	39	66
191	149	140	121	89	93	38	65
190	148	139	120	88	93	37	65
189	148	138	120	87	92	36	64
188	147	137	119	86	92	35	64
187	146	136	119	85	91	34	63
186	146	135	118	84	90	33	63
185	145	134	118	83	90	32	62
184	145	133	117	82	89	31	62
183	144	132	117	81	89	30	61
182	144	131	116	80	88	29	60
181	143	130	115	79	88	28	60
180	143	129	115	78	87	27	59
179	142	128	114	77	87	26	59
178	142	127	114	76	86	25	58
177	141	126	113	75	86	24	58
176	140	125	113	74	85	23	57
175	140	124	112	73	84	22	57
174	139	123	112	72	84	21	56
173	139	122	111	71	83	20	56
172	138	121	111	70	83	19	55
171	138	120	110	69	82	18	55
170	137	119	109	68	82	17	54
169	137	118	109	67	81	16	53
168	136	117	108	66	81	15	53
167	136	116	108	65	80	14	52
166	135	115	107	64	80	13	52
165	134	114	107	63	79	12	51
164	134	113	106	62	78	11	51
163	133	112	106	61	78	10	50
162	133	111	105	60	77	9	50
161	132	110	105	59	77	8	49
160	132	109	104	58	76	7	49
159	131	108	103	57	76	6	48

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 55-64

VERBAL				MANIPULACION			
Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
113	154	57	99	94	164	48	112
112	153	56	98	93	163	47	111
111	152	55	97	92	162	46	110
110	151	54	96	91	161	45	109
109	150	53	96	90	160	44	108
108	149	52	95	98	159	43	107
107	148	51	94	88	158	42	106
106	147	50	93	87	157	41	105
105	146	49	92	86	155	40	103
104	145	48	91	85	154	39	102
103	144	47	90	84	153	38	101
102	143	46	89	83	152	37	100
101	142	45	88	82	151	36	99
100	141	44	87	81	150	35	98
99	140	43	86	80	149	34	97
98	139	42	85	79	148	33	96
97	138	41	84	78	146	32	94
96	137	40	83	77	145	31	93
95	136	39	82	76	144	30	92
94	135	38	81	75	143	29	91
93	134	37	80	74	142	28	90
92	133	36	79	73	141	27	89
91	132	35	78	72	140	26	88
90	131	34	77	71	138	25	87
89	130	33	76	70	137	24	85
88	129	32	75	69	136	23	84
87	128	31	74	68	135	22	83
86	127	30	73	67	134	21	82
85	126	29	72	66	133	20	81
84	125	28	71	65	132	19	80
83	125	27	70	64	131	18	79
82	124	26	69	63	129	17	77
81	123	25	68	62	128	16	76
80	122	24	67	61	127	15	75
79	121	23	66	60	126	14	74
78	120	22	66	59	125	13	73
77	119	21	65	58	124	12	72
76	118	20	64	57	123	11	71
75	117	19	63	56	122	10	70
74	116	18	62	55	120	9	68
73	115	17	61	54	119	8	67
72	114	16	60	53	118	7	66
71	113	15	59	52	117	6	65
70	112	14	58	51	116	5	64
69	111	13	57	50	115	4	63
68	110	12	56	49	114		
67	109	11	55				
66	108	10	54				
65	107	9	53				
64	106	8	52				
63	105	7	51				
62	104	6	50				
61	103	5	49				
60	102	4	48				
59	101	3	47				
58	100	2	46				

Tabla 16. Sumas de las Puntuaciones a Escala y sus Equivalentes en Cocientes de Inteligencia

GRUPO DE EDAD 55 - 64

ESCALA TOTAL

Puntuación	C I	Puntuación	C I	Puntuación	C I	Puntuación	C I
		158	135	107	107	56	79
		157	134	106	106	55	78
207	161	156	133	105	106	54	78
206	161	155	133	104	105	53	77
205	160	154	132	103	105	52	77
204	160	153	132	102	104	51	76
203	159	152	131	101	103	50	76
202	158	151	131	100	103	49	75
201	158	150	130	99	102	48	75
200	157	149	130	98	102	47	74
199	157	148	129	97	101	46	74
198	156	147	129	96	101	45	73
197	156	146	128	95	100	44	72
196	155	145	127	94	100	43	72
195	155	144	127	93	99	42	71
194	154	143	126	92	99	41	71
193	154	142	126	91	98	40	70
192	153	141	125	90	98	39	70
191	152	140	125	89	97	38	69
190	152	139	124	88	96	37	69
189	151	138	124	87	96	36	68
188	151	137	123	86	95	35	68
187	150	136	123	85	95	34	67
186	150	135	122	84	94	33	67
185	149	134	121	83	94	32	66
184	149	133	121	82	93	31	65
183	148	132	120	81	93	30	65
182	148	131	120	80	92	29	64
181	147	130	119	79	92	28	64
180	146	129	119	78	91	27	63
179	146	128	118	77	90	26	63
178	145	127	118	76	90	25	62
177	145	126	117	75	89	24	62
176	144	125	117	74	89	23	61
175	144	124	116	73	88	22	61
174	143	123	115	72	88	21	60
173	143	122	115	71	87	20	59
172	142	121	114	70	87	19	59
171	142	120	114	69	86	18	58
170	141	119	113	68	86	17	58
169	140	118	113	67	85	16	57
168	140	117	112	66	84	15	57
167	139	116	112	65	84	14	56
166	139	115	111	64	83	13	56
165	138	114	111	63	83	12	55
164	138	113	110	62	82	11	55
163	137	112	109	61	82	10	54
162	137	111	109	60	81	9	53
161	136	110	108	59	81	8	53
160	136	109	108	58	80	7	52
159	135	108	107	57	80	6	52

APPENDIX II

RANDOM SAMPLING FROM TOTAL POPULATIONS
A DEMONSTRATION OF ITS POWER AND FLEXIBILITY

Russel F. Green

University of Rochester
and
University of Puerto Rico

APPENDIX II

RANDOM SAMPLING FROM TOTAL POPULATIONS A DEMONSTRATION OF ITS POWER AND FLEXIBILITY

Russel F. Green

University of Rochester
and
University of Puerto Rico

A Psychologist who gets involved in research must at some stage and at some level of consciousness decide how he is going to approach the problem of collecting samples of subjects on which to make observations or measurements of some kind. The usual solution is to use whatever subjects might be conveniently available and to hope for the best. There are, of course, exceptions to this solution but it is almost certainly the modal one.

The sampling problem is especially crucial when the project involves obtaining normative data for a population. In this context a serious effort to obtain a sample that is fairly and demonstrably representative of that population must be made. What is the best procedure to follow? The usual method is, I think, the so-called quota sampling method in which a central project directorate send instructions to regional offices that a certain number of people fitting certain combinations of control factors should be found and tested or observed in some manner. The regional offices then use various devices for finding such subjects and for persuading them to cooperate. Obviously, there are potentially important sources of bias in such methods but to what extent the biases actually occur is not known.

The sampling methods and results to be described here occurred as part of a project to translate into Spanish, to adapt to Spanish culture, and then to standardize the Wechsler Adult Intelligence Scale which is being carried out in Puerto Rico. The details of the design of the project have been described in some details elsewhere (1). Briefly, the goal is to produce an adult intelligence scale which has been adapted to Spanish culture in general rather than to the particular sub-culture of Puerto Rico. Some information as to the extent of cultural variation between Spanish speaking areas will also be obtained.

There are actually two reasons for reporting the results obtained by use of the random sampling methods adopted for the project.

First, along with illustrating the level of success of the methods, it should be more widely known that the resources for using similar methods are available in the United States. The resources are available for national, regional, state, or even smaller areas. Second, it is expected that from the work in Puerto Rico there will eventually be published a series of articles on matters that have been of considerable interest to Psychologists. The unique, for Psychology, sampling effort undertaken should make it possible to develop more nearly definitive answers to certain problems than have been developed in the past. The quality of answers proposed will rest very heavily on the sampling procedures reported here.

As an aside, perhaps it would be pertinent to merely mention now that in spite of real and significant differences that exist between the cultures of Puerto Rico and of the U.S. mainland, the Spanish language WAIS appears to have much the same statistical characteristics as the English language, i.e. U.S., WAIS. It does not follow, of course, that the adaptation of the WAIS has produced the best possible scale for Spanish culture. There are certain variations that would be interesting to try out. In the meantime this matter has simply been set aside and left unanswered.

In part by retention of materials from the WAIS and in part by new development, a set of Spanish language WAIS type materials was developed and pretested. A sample was then needed from which to carry out item and other analyses. Still later another, larger, sample was to be needed for the standardization. For the item analyses a representative sample of about 250 subjects was needed. The island population consisted at that time of about 2.6 million people. How should the sample be obtained? The decision was to attempt to use random methods. The methods used were such that they can be directly transferred to the United States although because of the size difference it would be considerably more expensive here.

The Department of Labor of the Island maintains a random sample of dwelling units for purposes of gathering and reporting labor statistics. Each sample drawn is subdivided into eight rotation groups and one of the eight is discarded each month. There are always at least two samples that are drawn and subdivided into such rotation groups and maintained by the Department of Labor. As rotation groups are dropped from one of the samples new ones are added from the next until, in time, each sample is used up.

When a rotation group is discarded it is made available to other government agencies or to research groups. The samples obtained for this project are being selected through a combination of systematic and random methods applied to the random set of dwelling units in the discarded rotation groups. In other words, a double randomization procedure is being employed.

An outline description of the Labor Department sampling methods will be given; this will be followed by a description of the methods used by the project; and finally the sample selected will be compared with the 1960 census and with the sample actually tested.

The Labor Department Sampling Methods

The Labor Department first identified five regions of the island, then divided these five regions into urban and rual zones, and finally added two special strata - housing projects started after 1960 and rural communities occupied after 1960. There were, then, a total of 12 strata. Urban was defined as towns and cities with 2,500 inhabitants or more. Rural was defined as 2,499 or less.

Next the 1960 enumeration districts (ED's) were arranged within each stratum. The ED's were arranged within each stratum as follows:

Urban

1. Cities and towns having 25,000 inhabitants or more.
2. Towns having 10,000 to 24,999
3. Towns having 2,500 to 9,999

Rural

1. Towns with 500 to 2,499 inhabitants
2. Open country

Towns in each bracket were arranged in geographical order. The census bureau had subdivided the three largest cities into tracts and these same tracts were also used by the Labor Department. Finally a serial number was assigned to each ED which indicated its order within its corresponding stratum.

It was necessary to develop an adjusted measure of size for each enumeration district. This was done by dividing the number of dwelling units by the average segment size. Average segment size was six dwellings in the Urban zones and 12 in the open country. The adjustment was to take account of the fact that some people live in group quarters (excluding inmates of institutions and the armed forces).

"Within each stratum a systematic random sample of enumeration districts was selected, with probability proportional to the measure of size. The sampling interval was seventy sets of six dwellings for urban and 12 sets for rural so that ED's with size larger than 70 segments could be selected more than once. Each selected ED is represented in the sample by one segment (of an average of six or 12 dwellings) for each time it was designated in the systematic selection."

"To make the selection, the measures of size of the ED's in each stratum were cumulatively added. Then for each stratum, a random number from one to 70 was selected to serve as a start. To this number, successive multiples of 70 were added to obtain the numbers corresponding to the selected ED's."

"After selecting the ED's these were divided at random into eight groups to be used in the rotation of the sample (as mentioned previously). Starting with a random digit from 1 to 8, successive numbers were assigned to the selected ED's, returning to 1 every time that 8 was reached. More than one digit was assigned to ED's which were selected more than once."

"The selection of the segments or clusters which will represent each ED in the sample is accomplished in two stages. First a 'block' or 'chunk' is selected (at random). A block is an area well defined by limits which are easily identified on the field, such as streets, roads, streams, etc. Then within the selected block a segment or cluster of (an average of six in Urban and 12 in Rural) dwelling units is selected (at random)."

There are other details concerning strata 11 and 12 and other matters involved in the process which need not concern us here. The above is the core of the Labor Department method of sample selection. It is not a simple random sample. A term to describe it would have to read something like the following: A region and zone controlled, three stage, area, systematic, random sample of dwelling units.

The Project Sampling Methods

The first item analysis sample was drawn from one of these Labor Department rotation groups. There being three systematic random stages in the process, it would seem that each of the final rotation groups should be acceptable as being essentially random sample of dwelling units.

The Labor Department is concerned only with people who are in or might enter the labor market. For this reason they make a census of people residing in the dwelling units but limit it to persons aged 14 and over. This census is also, of course, made available to users of the discarded rotation groups. One might wonder how precisely the "adult" population is sampled by a method in which population weights for enumeration districts are based on the total population. There seems, however, to be little cause for concern as will be illustrated a little later.

The first rotation group obtained from the Labor Department consisted of some 2471 persons aged 14 and over living in 900 dwelling units, 418 in the urban zone and 482 in the rural. The 418 urban units contained 1090 persons aged 14 and over (an average of 2.61 per unit) and the 482 rural units contained 1381 such persons (or 2.86 per unit).

The sample of persons to be used by the project was obtained from the Labor Department rotation group of dwellings in the following manner.[1] The residents in each of the housing units were listed, for most households, in the following order: (1) head of household; (2) spouse if any; (3) children 14 or over, if any; (4) relatives; (5) maids; and (6) others. The dwelling units were ordered by urban rural zone and municipality. The Labor Department was not entirely consistent in their labelling of occupants of dwelling units. Between this fact and the fact that there were varying numbers of persons in the dwelling units, no systematic effects from this ordering should have occurred.

Next a number from 1 to 5 was selected at random to determine the starting point. Starting with that subject, every fifth person was designated as a potential member of the sample. People selected were only potential members because some of them were under 16 and over 64 and also because it was decided not to use two blood relatives living in the same dwelling.

Due to a misunderstanding concerning the net number of subjects desired, too few people were obtained by the use of a selection interval of five. Four might have been a better interval to use but it would have increased considerably the number of blood relatives selected so perhaps this mistake was not really a mistake. In any event cases were added by selecting a new number between 1 and 15 (but excluding the number already used) at random and designating a second systematic subset.

These procedures produced a sub-sample of 617 from the Labor Department sample of 2471 persons in the 900 dwellings. The sub-sample averaged about .68 persons per dwelling.

The 617 cases were now divided in half by designating every odd one as being in sample 1 and every even one as being in sample 2. The stability (statistical reliability) of a sample of 617 was approximately maintained for each of these samples through the following procedure. The two halves of the 617 cases were compared on distributions of age, education, occupation, sex, zone, and region. Discrepancies between the two halves were largely balanced off by shifting some cases back and forth between them. When the control variables of the two halves were approximately balanced, a coin was tossed to determine which would be the primary sample and which would be the supplementary sample. The sampling error of the primary sample, while not estimated, should not have been much different from that of a sample of 617 cases.

The samples were used later in the following way. All persons under 15 and over 64 were dropped. Those who were listed as being 15 were kept in because it was expected that the time lag would be such as to result in most of them being 16 by the time project personnel contacted them. All but three of them were.

The individuals listed in the primary sample were then contacted and their cooperation solicited. Testers were given the name, address, sex, age, occupation, number of years of education, and zone of residence of each subject they were to contact. Some 82% of the subjects in the primary sample were tested. This meant that there were 18% losses and these were replaced from the supplementary sample. Replacements were made by selecting persons from the supplementary sample who were as similar as possible on all control variables to the lost cases.[1] At times this meant that one person might be replaced with another with, say, two years more education but when this happened another would be replaced with two years less. If it was impossible to make adjustments such as these, then any deviations that occurred were made in the direction of narrowing differences that might exist between the 1960 census data and the sample. This kind of event occurred in very few cases. It occurred primarily in relation to the Urban-Rural residence variable, and there in only about five or six cases.

It is obvious of course, that the procedures went beyond the present theory of sampling to provide known sampling distributions for the resulting means and variances of control variables. It is expected that because of the many stages of the procedure which incorporated random selection of sets of areas, dwellings, and finally people that the resulting sample can be regarded as essentially a simple random sample and statistical models based on this assumption applied to the data. To the extent thatthis is not so it seems likely that errors in estimating standard errors would be in the direction of making them too large. While errors of this kind are not necessarily desirable, they do seem to be more acceptable to more people than errors in the other direction.

Results

The essence of the procedures have been outlined. One can immediately spot various possible sources of bias. Was the sample of dwellings adequate? Does family size vary whith the kind as well as the location of the dwelling? In the latter case it was observed that almost 10% more people live in the average rural house than live in the average urban house? Was this adequately accounted for? This being a semi-tropical island, not all people live in houses! What effect might his have?

Even more important than these questions is that of whatever or not the replacement procedure was adequate to the task of preserving the characteristics of the original sample drawn. The original sample consisted of 303 persons age 14 and over. Of these, 42 were under 16 or over 64, which left 261. These 261 will be referred to as the primary sample. Forty three of the 261 were lost for various reasons. Now, three of the 43 replacements were also lost and re-placed. Hence 46 replacements were attempted. This means that efforts were made to contact and test a total of 261 plus 46 or 307 persons. Of these 307 persons, 254 or 82.7% were tested. Or, for those who would prefer it, of the original 261 persons 215 or 82.0% were tested.

The adequacy of the replacement procedures for the task of maintaining the characteristics of the original sample in the test sample, hinges to a considerable extent on the reasons why some 53 subjects were not tested. These reasons are summarized in Table 1. The really critical category in Table 1 is the refusals. Only 9 or an amazing 3% is all that refused. Two or 3 of the 6 dropped at the end might also have refused had they been contacted. Efforts to contact three of them had been made without success. Even so, the actual refusal rate was so low as to offer no threat of significant bias due to a willingness factor. Undoubtedly this low percentage was achieved in part because a three dollar honorarium was offered to each subject for his cooperation.

A word about the subjects dropped at the end should be said. These were dropped because of a combination of time schedules that needed to be met and because they were so scattered that it was proving very uneconomical to try to contact them.

Other categories seem self explanatory except perhaps for failures to find the subjects. In one case, for instance, when the tester went to the address, the house had been razed and the family had vanished. At times, and this may have been one of them, former neighbors proved unwilling to tell where someone lived or to where they had moved.

There are times in writing a report such as this that one feels strong urges to digress and to try to give something of the flavor and climate of the events going on. Recruiting, training, and keeping testers was at times a problem. Puerto Rican Spanish is typically slurred a great deal. A criterion adopted for gauging the adequacy of diction of the testers was whether or not the director of the project could understand! Some of the refusals of subjects were made by proxy. Husbands would sometimes refuse to allow wives to take the test. One hubsand beat up his wife for doing so without his permission. Two or three fathers ordered otherwise willing minor children not to take it. One wife ordered her husband not to! There were instances of long walks into remote hills and in some cases messages being sent in and the subjects walking out. And one could go on and on.

The purpose here, however, is to show that the methods worked. Ultimately, of course, the proof of the procedure has to lie in the question of how closely the sample, primarily the sample of people actually tested approximates what is known about the properties of the population. The following series of tables is intended to illustrate this.

Table 2 illustrates the degree of agreement that was achieved between the age and education distributions of the lost subjects and their replacements. The average education shifted only two tenths of a year and the average age only six tenths of a year. In both cases the figures compared are those contained in the Labor Department Files.

It should be pointed out that a few subjects did not report the same age and education to the project personnel as they had to the Labor Department. Some subjects admitted that they had reported themselves to the Labor Department as either being younger than they really were or as having had more schooling than they had had. They did this in the hopes that they might improve their chances of finding employment. It might actually be, therefore, that age and education characteristics of the population were more accurately e estimated by project data then by the Labor Department data.

Actually no method of gathering age information other than going to birth registers has proven accurate. This is illustrated in Table 3. Note that the census actually shows more people 35-39 than 30-34. The Labor Department statistics shows this even more and the sample tested also shows it. This is due to a long observed tendency on the part of people to understate certain ages, overstate others and to use convenient rounded numbers. People in their early forties with considerable frequency say they are "thirty nine" or younger.

The only shift of note between the sample tested and the census is that the sample shows a surplus of 2.2% in their forties and 2.3% in their fifties. This accurately reflects the increased longevity that has been achieved primarily over the last 25 years or so.

A two way classification comparing the education of the sample tested with the census and the primary sample in the urban and rural zones separately is presented in Table 4. For the most part the sample tested compares well with the primary sample. The principle exception to this is the zero education line. It proved surprisingly difficult to find some of the subjects with zero education and then when they were found they were not always testable. One was the idiot mentioned in Table 1. Others were mentally ill. Normally is they were testable they would cooperate. This difficulty was compensated for by adding subjects in the one and two year categories.

The primary sample differed from the census in certain ways. Definitely fewer persons with zero education appeared. It seems very likely that these differences are largely attributable to the fact that the average number of years of education is rising rapidly on the island and also to the fact that among the youth almost all now receive at least a few years of education. The differences, in other words, are in the direction of known population trends. The differences also are probably due in part to people's desire to qualify for employment.

Comparisons of the occupation by sex distributions of the sample tested with the census and the primary sample are presented in Table 5. The primary sample data for codes 92, 93, and 94 are subject to error in that this information was not supplied by the Labor Department. It was filled in later as the sample was contacted.

The category 94 requires some explanation. These represent people who were not in school, not in the labor force, and in the case of females not married. Eight per cent of males and ten per cent of females between 16 and 64 were so listed in the 1960 census. (Note the implications these figures have for unemployment statistics). These were pooled with the unemployed in the sample data because they could not be reliably differentiated. In the sample tested, however, six males and 10 females or 2.36 per cent and 3.94 per cent, respectively could be reliably described this way. The number of such people in the sample was definitely smaller than in the population. While this fits the population trend of greater employment opportunity, it is not clear that this much change occurred in four years.

Except for the shift in the 94 code, it is obvious that the samples drawn and tested fit the census quite well. There may be some slight trend for tested percentages to be higher than census percentages but this is by no means strong. The number of agricultural laborers is thought to be declining so a few, perhaps three or four, too many of these men may have been tested.

Comparisons of the sex and zone distributions of the sample tested with the census and the primary sample are presented in Table 6. Sex distributions were all remarkably similar. The zone distribution differed from the census by 3.2%. This difference probably is not statistically significant. Even if the sample is assumed to have the stability of a random sample of 617 a z-ratio for this difference of 3.2% would only be 1.6. The sample tested was closer to the census than to the primary sample.

Three way classification tables which include the joint distributions of sex x age x occupation, sex x region x zone, sex x age x zone, and sex x age x education have also been set up. Especially considering the fact that many N's in many subcells are very small, and in some cases zero, the comparisons in each are surprisingly good. Copies of them can be obtained from the author.

Summary

The sample differed from the 1960 census data in certain respects but in virtually every respect the differences were in the direction of known population trends. The percentage of people reporting zero education was lower than the census percentage, the percentage reporting high school attendance was higher, there was a lower proportion of non-labor force people in the sample reflecting improved employment on the island, people reported ages in anticipated patterns, but the number of students attending schools appearing in the sample may have been a little too small. Since differences were almost always in the expected direction it seems quite likely that the sample more accurately describes the population as it existed in 1964 than the 1960 census did.

As a post script, it should be pointed out that the procedure has now been replicated twice in Puerto Rico and once in Guatemala City, all with apparently similar outcomes. This leaves little doubt as to the essential accuracy of the method and procedure used.

The desirability of obtaining essentially random samples seem obvious. They leave one free to make quite positive inferences about population characteristics even when they are relatively quite small. This should be true regardless of the size of the population from which the sample is drawn. After all, it is the size and quality of the sample which determines the reliability of the inference. Population size may become important when greater size implies greater complexity and definitely becomes important when one is judging the practicality of carrying out such a procedure.

Footnotes

1. The actual sampling process was carried out by Dr. Miguel Valencia, statistician in the College of Commerce of the University of Puerto Rico.

2. Other methods of replacement might have been used, including just not replacing lost cases at all. However, the procedure adopted was intended to maintain the reliability of the sample which should have been approximately that of a sample of 617.

Table 1

REASONS FOR LOSSES FROM THE PRIMARY SAMPLE OF 261 AND FROM THE 46 SUBSTITUTES ATTEMPTED

*Refusals	9	
Moved to USA	6	
Could not be traced	10	
** Sick-Mental	5	
Sick-Physical	3	
Protocols Stolen!	2	
*** Test improperly administered	1	(not replaced)
Sibling or parent in Sample	5	
Spanish not native language	3	
Idiot (not testable)	1	
Dead	2	
Dropped at end of testing	6	
	53	

* The years of education of these 9 were 0, 0, 3, 9, 9, 11, 11, 11, 12

** Persons identified as mentally ill were excluded.

*** Of course, other mistakes were made in other test administrations mostly when digit symbol substitution was given to illiterates, but this was the only one dropped.

Table 2

COMPARISON OF AGE AND ED DISTRIBUTIONS OF LOST SUBJECTS AND THEIR REPLACEMENTS

AGE

	Lost	Replacements
60-64		
55-59	2	1
50-54	3	3
45-49	6	5
40-44	2	4
35-39	7	6
30-34	6	8
25-29	3	4
20-24	7	10
16-19	9	4
$\bar{X}$	32.8	32.2

EDUCATION

	Lost	Replacements
16+	2	2
15		1
14	1	1
13		
12	5	3
11	4	3
10		5
9	7	2
8	4	5
7	1	1
6	3	4
5	3	1
4	2	4
3	4	3
2	2	3
1	1	2
0	6	5
$\bar{X}$	6.9	6.7

Table 3

COMPARISON OF TESTED SAMPLE WITH CENSUS AND PRIMARILY SAMPLE ON AGE

Age	Census %	Primary Sample		Sample Tested		Census
	15-64	N	%	N	%	16-64
15	4.59	13	4.92			
16-19	15.61	45	17.05	43	16.93	16.34
20-24	14.01	33	12.50	40	15.75	14.67
25-29	11.14	25	9.47	23	9.06	11.66
30-34	10.33	24	9.09	23	9.06	10.81
35-39	10.68	30	11.36	31	12.20	11.18
40-44	8.72	15	5.68*	19	7.48	9.13
45-49	8.61	31	11.74**	21	8.27	9.01
50-54	6.08	18	6.82	24	9.45	6.37
55-59	5.40	15	5.68	15	5.91	5.65
60-64	4.71	15	5.68	15	5.91	4.93
		364		254		

* This total includes three subjects who were still 15 years of age when contacted.

** If considered alone, each of these percentages would be significantly different from those for the census (z = 2.5). The sum for the age span 40-49 is, however, very accurate.

Table 4

COMPARISON OF TESTED SAMPLE WITH CENSUS AND PRIMARY SAMPLE ON EDUCATION BY URBAN-RURAL ZONE

	URBAN					RURAL				
Years	Census	Primary		Tested		Census	Primary		Tested	
	%	N	%	N	%	%	N	%	N	%
16+	2.19	8	3.07	7	2.76	.34	0	.00	0	.00
14-15	2.13	5	1.92	6	2.36	.50	2	.77	2	.79
13	.98	2	.77	6	2.36	.33	0	.00	0	.00
12	5.44	21	8.05*	17	6.69	2.21	14	5.36*	14	5.51
10-11	4.34	15	5.75	20	7.87	2.58	15	5.75*	16	6.30
9	2.96	9	3.45	7	2.76	2.46	9	3.45	5	1.97
8	4.64	12	4.60	8	3.15	3.38	6	2.30	6	2.36
5-7	9.38	21	8.05	20	7.87	12.23	35	13.41	39	15.35
4	3.43	5	1.92	7	2.76	6.08	15	5.75	13	5.12
3	3.09	5	1.92	7	2.76	5.97	17	6.51	12	4.72
2	1.92	2	.77	3	1.18	3.33	9	3.45	14	5.51
1	1.04	3	1.15	6	2.36	1.87	5	1.92	6	2.36
0	5.76	8	3.07*	3	1.18	11.08	17	6.51*	10	3.94
No info.	.24	0	.00	0	0.00	.16	1	.38	0	.00
		116		117			145		137	

* Each star represents a deviation from the census that probably is significant. Every deviation is in the direction of known population trends.

Table 5

COMPARISONS OF THE OCCUPATIONS BY SEX DISTRIBUTIONS OF THE SAMPLE TESTED WITH THE CENSUS AND THE PRIMARY SAMPLE

Occupation	Code		Census % 16-64		Primary Sample 15-64*				Sample Tested 16-64			
			Male	Female	Male		Female		Male		Female	
I		Professionals			N	%	N	%	N	%	N	%
	00	General	1.56	.67	4	1.53	4	1.53	5	1.97	2	.79
	05	Teachers	.40	1.11	2	.77	0	.00	2	.79	1	.39
II	10	Farmers	1.29	.03	7	2.68	1	.38	5	1.97	0	.00
III	20	Managers	2.90	.44	7	2.68	5	1.92	7	2.76	4	1.57
	30	Office	1.77	1.95	3	1.15	6	2.30	1	.39	6	2.36
IV	40	Sales	2.36	.53	9	3.45	2	.77	8	3.15	1	.39
V	50	Craftsmen	5.05	.17	10	3.83	0	.00	14	5.51	0	.00
VI	60	Operators	5.20	2.91	17	6.51	7	2.68	20	7.87	5	1.97
	65	Clothing	.01	.44	0	.00	2	.77	0	.00	2	.79
VII	70	Domestics	.08	1.43	0	.00	6	2.30	0	.00	9	3.54
VIII	75-79	Services	2.24	1.38	8	3.07	5	1.92	11	4.33	3	1.18
IX	80	Ag. Labor	8.87	.15	19	7.28	3	1.15	24	9.45	1	.39
X	90	Laborers	2.87	.05	9	3.45	0	.00	7	2.76	0	.00
XII	92	Students	3.18	3.99	11	4.21	10	3.83	10	3.94	10	3.94
XIII	93	Housewives	--	25.23	-	--	75	28.74	-	--	71	27.95
	94	NS, NLF, NM	8.06	10.24	-		--	--	6	2.36	10	3.94
	98	Unemployed	2.12	.78	18	6.90	6	2.31	4	1.57	5	1.97
XI	99	No Info.	.29	.27	2	.76	3	1.14	0	.00	0	.00
					126		135		124		130	

* Three 15 year olds who were still 15 when contacted were omitted.

Table 6

COMPARISON OF TESTED SAMPLE WITH PRIMARY SAMPLE AND 1960 CENSUS-ZONE AND SEX

Zone	Census	Primary Sample		Tested Sample	
	%	N	%	N	%
Urban	47.6	116	44.4	117	46.1
Rural	52.4	145	55.6	137	53.9
		261		254	

Sex	Census	Primary Sample		Tested Sample	
	%	N	%	N	%
Male	48.7	126	48.3	124	48.8
Female	51.3	135	51.7	130	51.2
		261		254	